THE GRAMMAR OF MULTILINGUALISM

EDITED BY : Artemis Alexiadou and Terje Lohndal

frontiers Research Topics

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: **researchtopics@frontiersin.org**

THE GRAMMAR OF MULTILINGUALISM

Topic Editors:
Artemis Alexiadou, Humboldt University of Berlin & Centre for General Linguistics, Germany
Terje Lohndal, NTNU Norwegian University of Science and Technology & UiT The Arctic University of Norway, Norway

This volume investigates the nature of grammatical representations in speakers who master multiple languages. Since the early days of modern formal approaches to grammar, most work has been based on the language of monolingual humans. Less work has been conducted based on data from speakers who possess more than one language. Although important insights have been gained from a monolingual focus, there is every reason to believe that bi- and multilingual data can inform linguistic theory. A lot of ongoing work demonstrates that this is indeed the case, and the current volume contributes to this growing literature. Thus, the research topic addresses a number of questions relating to grammatical structures in multilingual speakers as well as the methodological issues that arise in the context of studying such speakers. A better understanding of the grammatical sides of multilingualism is crucial for understanding the human language capacity and in turn for offering better advice to the public concerning issues of language choice for multilingual children and adults, education, and language deficits in multilingual individuals.

Table of Contents

frontiers
in Psychology

Check for
updates

Editorial: The Grammar of Multilingualism

Artemis Alexiadou [1,2]* and Terje Lohndal [3,4]*

[1] *Institute of English and American Studies, Humboldt University of Berlin, Berlin, Germany,* [2] *Center for General Linguistics, Berlin, Germany,* [3] *Department of Language and Literature, NTNU Norwegian University of Science and Technology, Trondheim, Norway,* [4] *Department of Language and Culture, UiT The Arctic University of Norway, Tromsø, Norway*

The Editorial on the Research Topic

The Grammar of Multilingualism

Generative linguistics is primarily concerned with providing formal models of the linguistic competence of human beings. The goal is to adequately characterize and explain the structures of the grammar that each individual has constructed in his/her mind. This involves providing a formal description of the possible structures, which at the same time also rules out structures that do not occur. For example, a grammar of English should allow (1) but also rule out (2).

(1) John will eat cookies tomorrow.
(2) *John will cookies tomorrow eat.

The * is the indication that native speakers of English consider this sentence unacceptable. Differences between formal models need not concern us here; the important point that we want to make is that most formal models stay faithful to the following quote from Chomsky (1965, p. 3).

> "Linguistic theory is concerned primarily with an ideal speaker-listener, in a completely homogeneous speech-community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual performance."

Put differently, there has been an overwhelming monolingual focus within formal approaches to grammar. Although important insights have been gained from this focus, there is every reason to believe that a change of focus will prove very beneficial to formal models. More specifically, speakers who at some level of proficiency possess more than one language present a different set of data and theoretical challenges. We refer to all such speakers as "multilingual," well aware that the group is extremely heterogeneous. For present purposes, the exact breakdown of the group is not important, but major groups include individuals who grow up with multiple native languages, second and third language learners, and heritage speakers.

One of the groundbreaking aspects of generative linguistics has been to try to answer the question of what a possible mental grammar is. Specifically, the goal has been to unearth the structures that the human mind makes use of when it comes to language and at the same time develop theories and models that exclude those structures that do not seem to occur. From this perspective, data from multilingual speakers are essential since these speakers have grammars that often interact in ways that a theory of possible mental grammars needs to incorporate.

The current research topic addresses a number of questions relating to grammatical structures in multilingual speakers as well as the methodological issues that arise in the context of studying such speakers. The majority of the papers focuses on heritage speaker bilinguals. These are speakers who

are minority language speakers of a language acquired early on, which means that they are bilingual. Nevertheless, they are dominant in the majority language of the national community (see Montrul, 2008, 2016; Rothman, 2009 for much more). This leads to their characterization as unbalanced bilinguals. A typical trait of these speakers is that their grammar deviates in some way or other from the majority speakers of the relevant language. This makes it highly relevant to study which areas of the grammar are vulnerable and how this vulnerability should be understood: Is it because the acquisition of the heritage variety has been "incomplete" in some way, or is it because the grammar has attrited due to insufficient input? Some of these questions are explored in the current topic, highlighting a number of relevant factors that enter into our understanding of the nature of heritage grammars. Scontras et al'. review article focuses on the characterization of heritage speakers and what the study of these speakers can add to the study of linguistic competence. They offer a range of examples demonstrating their theoretical significance but also highlighting the methodological implications for the study of multilingualism more generally.

Corpora have become instrumental in the study of heritage speakers. Two papers contribute detailed studies of heritage speakers based on the same spoken corpus: The Corpus of American Norwegian Speech. Johannessen and Larsson study noun phrase-internal gender agreement and noun declension in a corpus of spoken American Norwegian. They argue that attrition affects agreement and not declension, and that complexity is an important factor in understanding the linguistic patterns. In the paper by Lohndal and Westergaard, gender in American Norwegian is explored further. It is shown that free-standing gender forms behave differently from suffixal declension class markers, and it is argued that transparency of gender assignment explains the vulnerability of the gender category.

Experimental methodology is pivotal in the study of multilingualism. Kim and Goodall present four formal acceptability experiments of island constructions in heritage Korean. They show that heritage speakers of Korean in the U.S. behave remarkably similar to native speakers residing in Korea, arguing that island phenomena are largely immune to environmental effects. Rather, island phenomena reveal deeper properties of the processor and/or grammar. Another experimental method is eye-tracking, which Arslan et al. use in a comparative study of how heritage speakers and late bilingual speakers of Turkish and German process grammatical evidentiality. They show that simplification takes place and they discuss how that should be interpreted theoretically.

Sometimes heritage speakers create new structures not seen in either of the two languages that are in contact. The paper by Yager et al. demonstrates exactly this point: They show that speakers of Heritage German have not simply lost dative case, rather, they have developed innovative structures to mark it, which are compatible with Universal Grammar. Again, we see the importance of studying various speaker and learner groups in order to get a better understanding of the kind of structures that the human mind is capable of generating.

Two of the papers in this research topic are concerned with language mixing in multilingual individuals. Chan considers mixing involving languages with contrasting head-complement orders, arguing that data from bilingual mixing or code-switching are highly relevant to better understand issues concerning phrase structure and linearization. Based on Persian-English bilinguals, Purmohammad conducts an experimental investigation of whether words from one of the bilingual speaker's languages can make use of the syntactic features from the other language, which he concludes is indeed possible.

Roeper is concerned with how to formally characterize the competence of multilingual speakers, notably second language speakers, arguing in favor of an approach based on Multiple Grammars. This approach holds that every speaker has a range of mental grammars, and Roeper presents numerous case-studies arguing in favor of this view. Rothman et al. are concerned with third language (L3) acquisition and how data from L3 speakers are theoretically important. They also show how L3 acquisition can benefit from employing neurolinguistics and psychological methodology to complement behavioral experiments.

Grohmann and Kambanaros are concerned with the role of language proximity, which is the closeness of the grammars that a child acquires, which they make use of to argue for an approach that they call "comparative bilingualism." Kaltsa et al. is a detailed study of coordinate subject-verb agreement in L1 and L2 Greek, showing that bilinguals behave similarly to monolinguals in terms of sensitivity to number agreement, although bilinguals are slower in processing overall. Lastly, Garraffa et al. consider linguistic and cognitive skills in Sardinian-Italian bilingual children, demonstrating significant similarity with monolinguals, although where there are differences, they are mostly in favor of bilingual children.

AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Montrul, S. (2008). *Incomplete Acquisition in Bilingualism: Re-examining the Age Factor*. Amsterdam: John Benjamins.

Montrul, S. (2016). *The Acquisition of Heritage Languages*. Cambridge, MA: Cambridge University Press.

Rothman, J. (2009). Understanding the nature and outcomes of early bilingualism:romance languages as heritage languages. *Int. J. Bilingualism* 11, 359–389. doi: 10.1177/1367006909339814

Heritage language and linguistic theory

Gregory Scontras[1]*, Zuzanna Fuchs[2] and Maria Polinsky[2]

[1] *Department of Psychology, Stanford University, Stanford, CA, USA,* [2] *Department of Linguistics, Harvard University, Cambridge, MA, USA*

This paper discusses a common reality in many cases of multilingualism: heritage speakers, or unbalanced bilinguals, simultaneous or sequential, who shifted early in childhood from one language (their heritage language) to their dominant language (the language of their speech community). To demonstrate the relevance of heritage linguistics to the study of linguistic competence more broadly defined, we present a series of case studies on heritage linguistics, documenting some of the deficits and abilities typical of heritage speakers, together with the broader theoretical questions they inform. We consider the reorganization of morphosyntactic feature systems, the reanalysis of atypical argument structure, the attrition of the syntax of relativization, and the simplification of scope interpretations; these phenomena implicate diverging trajectories and outcomes in the development of heritage speakers. The case studies also have practical and methodological implications for the study of multilingualism. We conclude by discussing more general concepts central to linguistic inquiry, in particular, complexity and native speaker competence.

Keywords: heritage linguistics, multilingualism, experimental methods, morphosyntax, syntax, semantics, pragmatics

INTRODUCTION

Since its inception, the generative tradition within linguistic theory has concerned itself primarily with monolingual speakers in its quest for what we know when we know (a) language. The object of study, linguistic competence, or grammar, instantiates in and emerges from the brains of human speakers. Grammar cannot get loaded onto a microscope slide or set upon a scale; it gets accessed through its effects on naturally-developing speakers who employ the grammar in their native language *du jour*. Grammar informs and determines linguistic behavior; linguists study grammar by studying the behavior of speakers and making generalizations about the idealized state of mind of these speakers. But which speakers?

The investigation of grammar is necessarily a circuitous enterprise: we observe linguistic competence through linguistic performance, the situation-specific deployment of grammar. But extra-linguistic factors influence performance, so linguists help themselves to various domain restrictions in an attempt to limit noise in the translation from competence to performance. Chomsky (1965, p. 4) provides an early description of the obstacle to be overcome: "The problem for the linguist, as well as for the child learning the language, is to determine from the data of performance the underlying system of rules that has been mastered by the speaker-hearer and that he puts to use in actual performance." Chomsky also provides an early characterization of one strategy for meeting this obstacle, focusing the linguist's attention on idealized, untainted language users:

Linguistic theory is concerned primarily with an ideal speaker-listener, in a completely homogeneous speech-community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual performance. (Chomsky, 1965, p. 3)

The rapid ascension of formal linguistics over the intervening five decades has demonstrated the success of this focused approach to the study of language (for a similar line of discussion, see Lohndal, 2013). A great deal of progress has been made to move beyond "grammars" in the traditional sense—comprehensive descriptions of language-specific regularities and their exceptions—to grammar in the Chomskyan sense: the rules and processes that generate those regularities in the first place.

Still, Chomsky's counsel necessarily excludes from study a wide swath of the world's language users, communities, and even languages. Put simply, the majority of speakers and speaking contexts fail to meet the admittedly idealized criteria above. But even ignoring the "grammatically irrelevant conditions" that govern the use of language, what do we make of the multitudes of speakers who may claim imperfect competence in more than one language? So far in the history of generative linguistics, the answer to this question has been "not much." Citing the wealth of data that gets ignored in such an unrealistic exclusion, together with the unique questions these data stand to answer, Benmamoun et al. (2013b, p. 129) propose we augment our study of language by "shifting linguistic attention from the model of a monolingual speaker to the model of a multilingual speaker." Similarly, Rothman and Treffers-Daller (2014) contend that multilingual speakers should be considered native in more than one language and call for a revision of the overall concept of a well-rounded native speaker. We follow these authors in focusing our attention on a subset of multilingual language users: heritage speakers.

To demonstrate the relevance of heritage linguistics to the study of language competence more broadly defined, this paper presents a series of in-depth case studies on heritage linguistics, documenting some of the deficits and abilities typical of heritage speakers. We adopt a modular approach to summarizing old and new findings, beginning with a look at the morphosyntax of agreement phenomena, then shift attention to the syntax of argument structure and of relativization; we then turn to the semantics and pragmatics of scope phenomena. The case studies we present serve double duty: first, their findings stand to characterize the similarities and differences between native and heritage speakers; and second, they engage with a popular strain of research in heritage language study, namely the various proposals meant to account for the near-native abilities of heritage speakers. Our aim is to show how the documented diversity of speaker profiles, abilities, and deficits requires a carefully nuanced approach to the study of multilingualism.

Before turning to the case studies, the remainder of this introduction describes the population of interest as it is typically characterized, together with various proposals meant to account for the unique linguistic competence of heritage speakers.

Introducing Heritage Speakers

To illustrate the defining characteristics of a heritage speaker, we begin with a few hypothetical examples. For starters, meet Samantha. Her family is from Korea, but she was born in Los Angeles and has never traveled to Korea. While in Los Angeles, Samantha grew up immersed in the rich Korean culture that is prevalent there (Los Angeles has the largest Korean-American population in the USA). Samantha went to a Korean Sunday school when she was a child, and she still uses Korean with her family and at church. However, she is more comfortable speaking in English; and although she reads Korean, she prefers reading in English. Samantha is always rather nervous about her Korean not being good enough for her family.

Margot is only a hundred or so miles south of Samantha, living in a secluded area in La Jolla, California (outside of San Diego). Her family moved there from Russia when she was three, and her younger siblings were all born in La Jolla. Her father still has some business in Russia, but Margot and her siblings rarely go there. They prefer traveling to Western Europe, where everybody speaks English and they have an easier time communicating. When Margot and her siblings meet other Russians, they are always a bit suspicious of them and do not socialize too much.

Doris grew up in a Jewish family in the Bronx. All her friends were Dominican and Puerto Rican immigrants; she still keeps in touch with some of them, and readily switches back and forth between English and Spanish when they chat. Doris took Spanish in high school and quickly discovered that the language she learned from her friends was vastly different from the language in her textbook; she recalls the experience in her Spanish class as a nightmare. "Every time I spoke, my teacher mocked and belittled me for saying everything wrong. Apparently what was right for my friends was not right for the Anglo woman who was teaching me…"

Robert was born in Frankfurt, but when he was just a few months old, his family moved to Abu Dhabi, where his father worked as a banker. He had an Arabic-speaking nanny and went to an international school, but socialized with Arabic-speaking children (they all shared a passion in soccer). Robert moved back to Germany when he was 15, got his education in Germany, and is currently living in Berlin where he works as a graphic designer. He is still in touch with his friends in Abu Dhabi—they connect over social media—and it is his hope to save enough money to travel back to the place where he spent his childhood.

Shawn was born in Canada. His mother is Japanese and his father is British, fluent in Japanese. The family moved to Japan when Shawn was a toddler. He has received all of his education in Japanese, and although he has had a fair amount of English instruction and speaks English with his father now, as a young adult, he is more comfortable in Japanese. Recently, he took a course in American literature in his college; whenever possible, he tried to read the assigned books in a Japanese translation, which he found much easier than the original English.

What do these people have in common? They were all exposed to a certain language in their childhood, but then switched to

another language, the dominant language of their society, later in their childhood. These are unbalanced bilinguals, sequential (Doris and Margot) or simultaneous (Robert, Shawn, Samantha), whose home language is much less present in their linguistic repertoire than the dominant language of their society. They may have gotten there in different ways, but they are all heritage speakers.

Narrowly defined, heritage speakers are individuals who were raised in homes where a language other than the dominant community language was spoken, resulting in some degree of bilingualism in the heritage language and the dominant language (Valdés, 2000). A heritage speaker may also be the child of an immigrant family who abruptly shifted from her first language to the dominant language of her new community. Crucially, the heritage speaker began learning the heritage language before, or concurrently with, the language which would become the stronger language. That bilingualism may be imbalanced, even heavily imbalanced, in favor of the dominant language, but some abilities in the heritage language persist.

Heritage speakers present a unique testbed for issues of acquisition, maintenance, and transfer within linguistic theory. In contrast to the traditional acquisition trajectory of idealized monolinguals, heritage speakers do not seem to exhibit native-like mastery of their first language in adulthood. As the definition of the heritage speaker makes clear, this apparent near-native acquisition owes to a shift of the learner's attention during childhood to a different dominant/majority language. However, the specifics of this attainment trajectory are anything but clear.

Developmental Trajectories of Heritage Speakers

The pathways to heritage speakerhood vary quite widely. Similarly diverse is the range of abilities that result. It should come as no surprise, then, that the proposed trajectories to the competence of heritage speakers are at least as complex as the speakers and abilities they are meant to characterize. Here we consider possible outcomes in the shape of heritage grammars. Setting aside the possibility that the heritage grammar can match that of the native baseline (something that we do not discuss in this paper, if only for lack of space), at least three other outcomes are possible: *transfer* from another grammar, *divergent attainment*, and *attrition* over the lifespan. Crucially, behavior with different grammatical phenomena may derive from diverging outcomes, owing in part to the broader linguistic context. Ultimately, research in heritage languages should be able to predict a particular outcome for a given phenomenon or context, but the field is not there yet. For now it suffices to survey the possibilities.

Types of Outcomes
Dominant language transfer
An important point of contact between heritage speakers and second language learners lacking from traditional L1 acquisition is the interplay between the learner's first (heritage) language and second (dominant) language. Language transfer, or the nature of that particular interplay, is a foundational issue in second language acquisition research: to what extent does the

first language grammar play a role in shaping the developing second language grammar? The effects of the native language on the acquisition of a second language in different levels of linguistic analysis (e.g., phonology, morphology, syntax, semantics, or the lexicon) have been extensively documented in the second language acquisition literature (e.g., Odlin, 1989; White, 1989; Gass and Selinker, 1992; Schwartz and Sprouse, 1996; Jarvis, 1998). The question of transfer arises in other language contact situations, including pidgin and creole genesis, where phenomena like lexical borrowings and so-called "areal features" are the well-known consequences of language contact. Research on bilingualism and language contact also suggests that the direction can reverse, such that the second language encroaches on the structure of the native language in systematic ways (Seliger, 1996; Pavlenko and Jarvis, 2002; Cook, 2003).

With the knowledge that grammar is a porous vessel whose contents are susceptible to contamination, in examining the linguistic characteristics of heritage grammars, the first question that often comes to mind is whether many of the "simplified," non-standard characteristics observed in the heritage grammar could be due to transfer from the dominant language. For example, one can readily entertain the possibility that nominal and verbal inflectional morphology in Spanish and Russian heritage speakers gets eroded because the contact language in most of the heritage speakers tested to date is English, a language which does not mark gender on nouns or have rich tense/aspect and mood morphology. The same explanation goes for the preference for SVO word order over topicalization, which in turn leads to greater word order rigidity.

An obvious way to resolve this question over the source of simplified characteristics in heritage grammars is by testing heritage speakers whose majority language is typologically close to their heritage language (Spanish heritage speakers in Italy or Brazil, for example); ensuring that the contact language is at least as complex as the target language with respect to the phenomenon of interest controls for possible simplification transfer. Another option is to isolate the effects of different contact languages, either by comparing the effects of different dominant languages on one and the same heritage language, or by comparing the effect of one and the same dominant language on different heritage languages. In either case, one must take care to determine the status of the phenomenon of interest in *both* the heritage *and* the dominant grammar, to see whether there is anything to transfer in the first place. Put differently, comparison with a native speaker baseline does not suffice to prove transfer, as the native baseline might differ in important ways from its manifestation in the heritage population. We return to this cautionary tale below, and in our fourth case study, on scope calculations.

Divergent attainment
Heritage speakers are early bilinguals who learned their second (majority) language in childhood, either simultaneously with the heritage language, or after a short period of predominant exposure to and use of the minority language. A common pattern in simultaneous bilinguals is that as the child begins to socialize in the majority language, the amount of input from and use

in the minority language is reduced. Consequently, the child's competence in the heritage language begins to lag, such that the heritage language becomes, structurally and functionally, the weaker language. Developmental delays that start in childhood never eventually catch up, and as the heritage child becomes an adult, the eventual adult grammar does not reach native-like development. This trajectory was originally introduced in the literature as "incomplete acquisition" (Polinsky, 2006; Polinsky and Kagan, 2007; Montrul, 2008; Benmamoun et al., 2013b); however, some researchers have argued against the use of this term because it has negative connotations (e.g., Pascual y Cabo and Rothman, 2012) or covers arguably unrelated phenomena, namely lack of mastery due to limited input vs. lack of knowledge associated with education and exposure to a standard dialect (e.g., Pires and Rothman, 2009). In this paper, we will be referring to the phenomenon as "divergent attainment," in hopes that this term is more agreeable. Moving beyond the terminology, it is crucial to focus on contexts where such an outcome can be predicted; this is one of the larger goals of heritage language research.

A clear example of divergent attainment is the acquisition of the subjunctive in Spanish. Blake (1983) tested monolingual children in Mexico between the ages of 4 and 12 on their use of the subjunctive. He found that between the ages of 5 and 8, knowledge and use of the subjunctive was in fluctuation; children did not show categorical knowledge of the Spanish subjunctive until after age 10. Heritage speakers who received less input at an earlier age and no schooling in the language never fully acquire all of the uses and semantic nuances of the subjunctive, as reported in many studies (Silva-Corvalán, 1994; Martínez Mira, 2009; Montrul, 2009; Potowski et al., 2009; see also Silva-Corvalán, 2003, 2014, for longitudinal observations). It would seem, then, that the subjunctive employed by adult heritage speakers of Spanish evidences a calcified version of its attainment in monolingual youth.

Attrition
Distinct from, but not mutually exclusive with attainment is the outcome of attrition. Under normal circumstances, L1 attrition refers to the loss of linguistic skills in a bilingual environment. It implies that a given grammatical structure reached full mastery before suffering weakening or being subsequently lost after several years of reduced input or disuse. Thus, attrition is "the temporary or permanent loss of language ability as reflected in a speaker's performance or in his or her inability to make grammaticality judgments that would be consistent with native speaker monolinguals of the same age and stage of language development" (Seliger, 1996, p. 616). Attrition over the lifespan is a particularly intriguing case, since it challenges the common assumptions concerning the stability of structural change in adults.

Attrition often occurs during the first generation of immigration, affecting structural aspects of the L1 due either to language shift or to a change in the relative use of the L1 (De Bot, 1990)[1]. Attrition can also occur much earlier, having more

dramatic effects on the integrity of the grammar. Recent research suggests that the extent of attrition is inversely related to the age of onset of bilingualism (Pallier, 2007; Montrul, 2008; Bylund, 2009; Flores, 2010, 2012). Prepubescent children tend to lose their L1 skills more quickly and to a greater extent than people who moved as adults and whose L1 was fully developed upon migration (Ammerlaan, 1996; Hulsen, 2000). That is, the extent of attrition and severe language loss is more pronounced in children younger than 10 or 12 years old than in individuals who immigrated after puberty. Research has also shown that severed or interrupted input in childhood, as in international adoptees, leads to severe attrition, including total language loss (Montrul, 2011).

There are two ways to tease apart divergent attainment and attrition in later childhood. The first strategy consists of conducting longitudinal or semi-longitudinal studies of children, like the ones by Anderson (1999), Merino (1983), and Silva-Corvalán (2003, 2014). These authors were able to document the incremental accumulation of errors in agreement (i.e., case or gender marking) in their investigation of immigrant children who arrived in their new country around age 8;0 or older. Their results show a significant accumulation of errors, which eventually leads to the loss of a baseline pattern. Still, it has yet to be determined at what point such error accumulation reaches the point of no return, resulting in severe language loss.

The other strategy for teasing apart attrition and divergent attainment compares children and adult heritage speakers. If it can be shown that normally-developing child heritage speakers perform better than their adult counterparts, then we have evidence for attrition. This strategy serves as the basis of our second and third case studies, which compare heritage speakers with monolingual controls, as well as with monolingual and heritage children.

What Motivates the Outcomes?
Having suggested three possible ways in which heritage language may differ from the baseline, we turn next to the potential sources for such differential outcomes. We explore three different scenarios: changes in the input, general constraints on memory, and universal structural principles.

Incipient changes in the input
To understand the source of seemingly non-native abilities in heritage language speakers, we must establish whether the immigrant communities themselves speak an altogether different variety from that spoken in the country where the language is dominant. In other words, it is important to ascertain patterns of language maintenance or change in the variety used by the immigrant community, to determine the input heritage language learners are receiving. Thus, one ought to determine whether the first generation grammar shows any of the non-standard properties attested in the heritage language; this approach is typical of sociolinguistic studies (Otheguy and Zentella, 2012). If the first generation grammar already shows signs of drift from

---

[1]Until recently, the vast majority of studies on language attrition were conducted with elderly adults (Levine, 2001; Schmid, 2011), who attained full linguistic

competence before attrition began and who may also show independent aging effects.

the standard baseline, then the culprit is not the heritage learner. Conversely, if a property is not part of the register spoken to the heritage speakers, then it cannot be acquired, but must be the result of reanalysis or innovation.

To see the value in considering the grammar of first-generation immigrants in the shaping of heritage grammar, consider the findings of Montrul and Sánchez-Walker (2013), who tested differential object marking (DOM) in English-dominant heritage speakers of Spanish, first-generation immigrants (the input to the heritage speakers), as well as L1 speakers of different age cohorts in Mexico. The authors found that the child and adult heritage speakers omitted DOM, but so did the first-generation immigrants. The question then becomes: why did the input change in the first place? Answering this question brings us to two additional sources for the divergence between native and heritage grammars: general resource constraints (e.g., memory constraints) becoming more pronounced in a less dominant language, and universal structural properties of grammar extending their influence.

Resource constraints

Some changes in heritage language consist of constraining the domain within which a particular property applies. A recent example of this type of finding comes from Kim's (2007) study of binding interpretations by Korean heritage speakers in the USA and China. The study tested knowledge of binding interpretations with local and long-distance anaphors. Here we see deployed one of the suggestions made earlier for isolating the quality of transfer from a dominant language: comparing the effects of different dominant languages on one and the same heritage language. In many respects, Chinese and Korean are more similar than Korean and English. As such, Korean heritage speakers in China, who suffered less interference from their dominant language, were expected to be more accurate with long-distance binding than the Korean heritage speakers in the USA. However, Kim found that the two groups of Korean heritage speakers still had a marked preference for local binding, regardless of the contact language. Thus, the result state— loss of long-distance binding in heritage Korean—appears to have derived not from contact with a specific different system, but from contact with *any* different system. In other words, once the heritage language loses ground to another dominant language, whichever that language might be, resource-intensive phenomena like binding (or scope inversion; see Section *At the Interface: Scope Interpretations*) become more restricted.

The loss of long-distance binding in heritage Korean appears to be an instance of general constraints on memory becoming more pronounced in heritage speakers: shorter dependencies are preferred because they make fewer demands on the parser's memory. Given that the heritage speaker is already performing the costly task of speaking in a less dominant language, the cost of resource-intensive operations explodes, sometimes to the point of totally obscuring the availability of the operation.

Universal principles of language structure

In heritage grammars, where speakers are limited in their deployment of complex grammatical phenomena, language structure sometimes follows what looks like a default design, employing a seemingly restricted set of grammatical categories and operations. The list of default-like structures attested for heritage languages includes the use of dependencies which target only the highest structural constituent (as in the Russian relativization discussed in Section *Relativization: In Support of Universal Structural Principles*); the absence of nesting dependencies (Benmamoun et al., 2013a,b); the elimination of irregular morphology and the concomitant rise of analyticity (Benmamoun et al., 2013a,b); rigid word order (Isurin and Ivanova-Sullivan, 2008; Ivanova-Sullivan, 2014), often accompanied by the placement of closely associated items next to each other, in keeping with Behaghel's First Law (Behaghel, 1909; Haiman, 1983); and the lack of non-compositional structures (Dubinina, 2012; Rakhilina and Marushkina, 2014). All of these properties appear to at least superficially make the heritage language more user-friendly, in accord with general properties of language structure.

However incomplete, this list of properties bears a striking similarity to recurring traits observed in creole languages and often associated with the underlying innate principles of language structure, as in Bickerton's famous Bioprogram (Bickerton, 1984, 1988). We are not trying to propose a new version of the Bioprogram here, but we would like to offer two considerations. The first one is obvious: since there appear to be recurrent features observed in heritage language, a comprehensive list of heritage-language-specific properties related to universal principles of *optimal* language design is needed. Such a list needs to be established empirically, on the basis of a larger set of studies, and then re-evaluated in light of linguistic theory. Doing so would allow us to understand in a more coherent way the notion of language defaults and optima. Relatedly, given the initial evidence for their reliance on universal language principles, heritage speakers have a great deal to offer linguistic theory, because they speak directly to Plato's problem in language: showing how a grammar can be acquired under conditions of reduced input and usage. This reality makes heritage languages a desirable object of investigation, and we need to learn how to use them better to enrich the debate about the nature of the language faculty.

This completes our brief introduction to the population we herewith study: heritage language speakers. A reader interested in more details of this group can find further discussion in Benmamoun et al. (2013a,b), Montrul (2008) and Polinsky and Kagan (2007). In the remainder of this paper, we examine in considerable detail specific properties of heritage language grammar through a series of case studies. In doing so, we pursue two interconnected goals. First, we present theoretically relevant phenomena whose status in heritage language serves as evidence for a particular trajectory or outcome, either contrasting with the native baseline (as with morphosyntax in Section *Agreement Morphology and Category Structuring*) or in support of general structural principles (as with syntax in Sections *Argument Structure: The Unaccusative Challenge* and *Relativization: In Support of Universal Structural Principles*). Second, by concentrating on areas of known vulnerability in language structure, we show

that the ultimate fate of vulnerable domains can vary depending on the level or type of representation and its specific language context.

We begin our investigation with a look at morphosyntax, agreement in particular (Section *Agreement Morphology and Category Structuring*). We then analyze phenomena related to argument structure (Section *Argument Structure: The Unaccusative Challenge*) and syntactic dependencies (Section *Relativization: In Support of Universal Structural Principles*). In Section *At the Interface: Scope Interpretations*, we venture outside narrow syntax and consider the grammar of scope, which brings together several interfacing grammatical domains. Section *Conclusions* presents our conclusions, where we revisit the question of what it means to be a native speaker, and what linguists stand to gain from embracing the reality of heritage linguistics.

AGREEMENT MORPHOLOGY AND CATEGORY STRUCTURING

In our first case study, we extend previous work on the morphosyntax of agreement in Spanish. Given the well-documented difficulty heritage speakers display with morphology in general and agreement morphology in particular (see Benmamoun et al., 2013b, pp. 141–144, and further references therein), we expected to find differences between native and heritage speakers of Spanish, and, more importantly, we expected these differences to be informative with respect to the agreement mechanism and its features in these minimally-differing grammars. But before asking how heritage speakers of Spanish perform, we must first establish the native baseline.

In Fuchs et al. (in press), we investigated the organization of number and gender features in Spanish, bringing experimental evidence to bear on the structure and content of agreement. The choice of number and gender features was not accidental: the third class of agreement features, person, stands apart both descriptively (for example, unlike the other features, person agreement never appears on adjectives; see Baker, 2008) and theoretically (cf. the hierarchical positioning of person in the feature geometry of Harley and Ritter, 2002). Meanwhile, the relationship between gender and number is less clear. Assuming that both features are represented in syntax, there are two analytical possibilities, both proposed in the literature. According to one scenario, gender and number are always bundled together (cf. Ritter, 1993; Carstens, 2000, 2003). Under the bundling model, number and gender features are projected and valued together; the valuation of gender presupposes a valuation of number, as gender features do not project independently of number. The bundling model draws its empirical inspiration from the fact that languages regularly combine gender and number information in the morphology; one rarely finds systems where the two features participate in agreement and yet are independent of each other.

In the alternative, split model (Picallo, 1991; Antón-Méndez et al., 2002; Carminati, 2005), gender morphology hosted on a nominal stem heads its own syntactic projection (GenP),

and GenP is dominated by NumP (i.e., the source of number features/morphology). Thus, number and gender features are projected—and therefore also valued—independently of each other. One of the major arguments in favor of the split model comes from the order of morphemes in nominal derivations. In those languages where number and gender morphology can be descriptively separated, the order is Stem-Gender-Number, as in the following Spanish examples:

(1)   a. $[[libr]\text{-}[_{GenP}\ o\text{-}]\ [_{NumP}\ s]]$ 'books'
       b. $[[libr]\text{-}[_{GenP}\ o\text{-}]\ [_{NumP}\ \emptyset]]$ 'book'

Because it levels the hierarchical distinction between number and gender, the bundling model does not have a straightforward way of predicting the ordering in (1). That the split model derives such an order is a side effect of the simple feature geometry: number dominates gender[2]. But which model, bundling or split, is the right one for Spanish? This was the question we set out to answer in Fuchs et al. (in press).

In Spanish, number and gender are expressed through independent suffixes. For gender, the word marker -*a* most often corresponds to the feminine, and the word marker -*o* most often corresponds to the masculine (although see Harris, 1991, for a more detailed discussion and many exceptions). Number is represented much like it is in English: The plural is marked by -*s*, whereas the singular receives no marking. Determiners and adjectives must agree with the noun in both number and gender.

(2)   a.   la   manzana   b.   el   plátano
             the.F.SG   apple.F.SG   the.M.SG   banana.M.SG
       c.   las   manzanas   d.   los   plátanos
             the.F.PL   apple.F.PL   the.M.PL   banana.M.PL

As the number and gender agreement morphemes are in principle independent, we could manipulate their combination to produce sentences with different kinds of agreement errors in the Fuchs et al. study. Because the bundling and split models of feature geometry make different commitments regarding the valuation of agreement features, the predictions of the two models pull apart in cases of agreement attraction. In such cases, like the English example in (3), a noun (italicized) intervenes between the head noun (underlined) and its predicate (in bold), and the predicate incorrectly enters into agreement with the intervening noun rather than the head noun (in (3), *were* is plural, but should be singular to match the number of the head noun *key*). Because features of the local noun match features of the predicate, people incorrectly perceive the sentence as grammatical. This is agreement attraction.

(3)   The <u>key</u> to the *cabinets* **were** lost.

Cases of agreement attraction have been experimentally studied in various languages, testing whether there is an asymmetry between different values of features in triggering agreement

---

[2]For other considerations, both empirical and theoretical, that have gone into the debate about bundling vs. split models, see Alexiadou (2004), Kramer (2014), and Ritter (1993).

errors (e.g., English: Bock and Miller, 1991; Bock and Eberhard, 1993; Vigliocco et al., 1996; Vigliocco and Nicol, 1998; Bock et al., 2012; Spanish: Vigliocco et al., 1996; Antón-Méndez, 1999; Antón-Méndez et al., 2002; Alcocer and Phillips, 2009; Lago et al., 2015; Italian: Vigliocco et al., 1995; Vigliocco and Franck, 1999; French: Vigliocco et al., 1996; Dutch: Bock et al., 2001; Dutch and German: Hartsuiker et al., 2003; Russian: Lorimor et al., 2008). In Fuchs et al., we extended the method by putting the phenomenon of attraction to use in exploring the difference between bundling and split approaches.

Recall that if number and gender are bundled, then they ought to be valued simultaneously. This suggests that the number and gender features of a noun should determine agreement together, at the same time. When an incorrect noun enters into agreement with an adjective, both its number and gender features should effect agreement attraction. To illustrate this point, consider the following ungrammatical sentences:

(4) a. *El   niño   considera   la   noticia
         the.M.SG   boy   consider.PRES.3SG   the.F.SG   news.item.F.SG
         en los   periódicos   terriblemente   aburridos.
         in the.M.PL   magazine.M.PL terribly   boring.M.PL
         ('The boy considers the news item in the magazines to be terribly boring.')

     b. *El   niño   considera   la   noticia
         the.M.SG   boy   consider.PRES.3SG   the.F.SG   news.item.F.SG
         en   las   revistas   terriblemente   aburridas.
         in   the.F.PL   magazine.F.PL   terribly   boring.F.PL
         ('The boy considers the news item in the magazines to be terribly boring.')

Both (4a) and (4b) are ungrammatical. However, in each sentence the local noun has entered into agreement with the adjective, which may lead to an illusion of grammaticality via attraction. If number and gender are projected and valued together, per bundling approaches, then when the probe (incorrectly) gets a feature (e.g., number) from the local noun, it should be able to get the other feature (e.g., gender) as well. In other words, agreement attraction in one feature ought to precipitate agreement attraction in the other feature, with the result that both of the above sentences should be rated equally high (or equally low).

If, however, number and gender are split, then they are projected and valued independently, and agreement attraction in number can proceed independently of agreement attraction in gender. This means that, all other factors being equal, a violation in gender agreement may be judged higher or lower than a violation in number agreement. Crucially, the violations are evaluated on their own merits. Furthermore, if the two features are independent of each other, we can expect that a violation in both of them would be more offensive to a comprehender than a violation in just one feature. This expectation is based on the observation that the more grammatical constraints violated, the higher the degree of degradation (consider Kluender, 2004). Applying that logic, we expected that the violation in (4a), where both the gender and the number of the head noun are mismatched, should be rated lower than (4b), where only the number feature is mismatched. Thus, under a split model, (4a) should receive a lower rating.

We originally tested native speakers of Spanish (n = 50) in an auditory sentence-acceptability rating task involving sentences as in (4), with differing numbers of agreement errors. In each of these critical conditions, the head noun appeared in the singular while the local noun and adjective appeared in the plural. By permuting the gender of the head noun, the local noun, and the adjective, we engineered potential attraction conditions in which the local noun either agreed with the adjective in only number (i.e., both were plural, but their gender did not match), or in both number and gender. Participants heard a recording of the sentence, and then were asked to rate its acceptability on a 5-point Likert scale (1 = "completely unacceptable"; 5 = "completely acceptable"). The results are plotted in **Figure 1**, which organizes ratings by potential attraction condition; error bars represent bootstrapped 95% confidence intervals drawn from 10,000 samples of the data.

For feminine head nouns, the sequence with a single agreement error, **F.SG**—F.PL—F.PL, was rated significantly higher than the sequence with two agreement errors, **F.SG**—M.PL—M.PL[3]. Thus, we found evidence of attraction such that ungrammatical sequences were accepted, but attraction occurred only between the number features of the local noun and adjective; if the gender of the head noun did not match that of the adjective, the sentence was correctly viewed as sub-par. For masculine head nouns, the difference between ratings given for single-error attraction conditions (**M.SG**—M.PL—M.PL) and double-error attraction conditions (**M.SG**—F.PL—F.PL) was not significant; we failed to find evidence of attraction at all for masculine head nouns.

Given the predictions of the bundling vs. split models, we interpreted the asymmetry in the ratings of agreement mismatches for feminine head nouns as evidence that number and gender features are valued separately; were they valued together, we should have found no difference between the conditions in which only one feature determined attraction effects and the conditions where both features caused attraction. Thus, in Spanish, a split model of number and gender features best accounts for the data: these features are treated separately in agreement.

Now, given the precarious status of agreement morphology in heritage grammars, our question shifts to whether heritage speakers diverge from native ones in their agreement behavior,

---

[3]Here and below, the gender/number of the head noun appears first, in boldface.
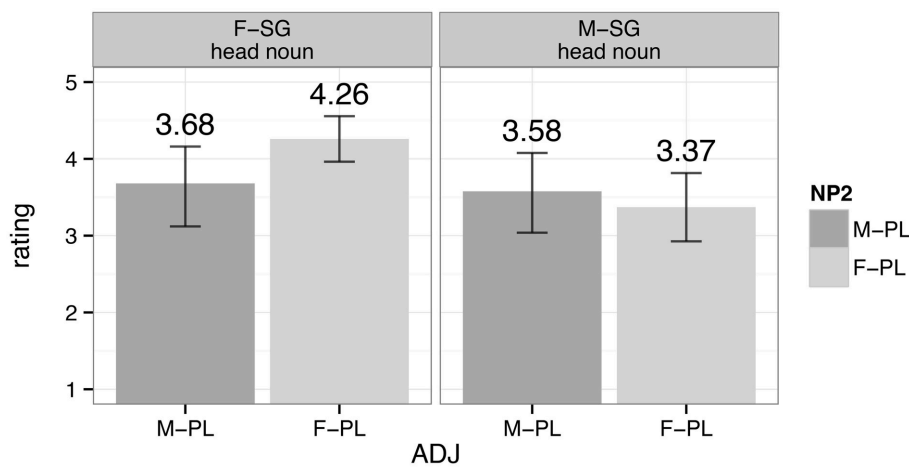
FIGURE 1 | Average ratings for potential attraction conditions (Fuchs et al., in press).

such that their representation of number and gender features is fundamentally different from the baseline. We extended the auditory sentence-acceptability rating task from Fuchs et al. to English-dominant heritage speakers of Spanish, as well as baseline controls. The results appear in **Figure 2**.

Note first that the results of our new population of native speaker controls ($n = 28$) replicate those found in the original study: participants perceived conditions with agreement errors in both number and gender as ungrammatical and rated them lower than conditions with an agreement error in only one feature; and feminine vs. masculine head nouns were treated differently.

Turning to heritage Spanish, we identified these speakers on the basis of a demographics questionnaire that preceded testing. Heritage speakers ($n = 71$) were those who indicated that they first learned Spanish and then English, had no formal education in Spanish, and who never lived in a Spanish-speaking country during childhood. **Figure 2** shows that heritage speakers behave similarly to the native baseline in treating feminine vs. masculine head nouns separately with respect to attraction. However, unlike native speakers, heritage speakers rated attraction conditions equally high, regardless of the number of agreement mismatches between the head noun and the adjective. As long as the attractor noun agreed with the adjective in at least one feature, attraction succeeded and participants rated these ungrammatical sentences as acceptable.

The most straightforward interpretation of these results, in accordance with our original predictions for the native baseline, would have heritage speakers bundle number and gender features so that they are projected and valued together. However, before jumping to this conclusion, we must be realistic about the morphological limitations in heritage language, limitations that motivated the current study in the first place. What if the observed insensitivity to the number of agreement errors signaled not that number carries gender along for the ride while it gets valued in the heritage grammar, but rather that our heritage participants did not access gender as they processed the data presented to them? In other words, it could be the case that our

heritage speakers simply ignored gender altogether. While we lack conclusive evidence to tease apart bundling from ignorance (i.e., from the ignoring of gender), the differential treatment of feminine vs. masculine head nouns in accord with the native baseline suggests that at least at some level, heritage speakers are attending to gender. If we take this evidence seriously, then heritage speakers have reanalyzed the feature system of Spanish so that it levels the hierarchical distinction between number and gender. Put simply, what native speakers treat as separate categories (i.e., number and gender), heritage speakers handle as but one, thus opting for the bundling of these categories. The result is a different, ostensibly simpler grammar than that of the baseline.

ARGUMENT STRUCTURE: THE UNACCUSATIVE CHALLENGE

Having considered differences in the domain of morphosyntax, we now leave the "morpho" component behind and dive head-first into syntax. But which syntactic phenomena might undergo change in heritage languages? Atypical, complex, or infrequent constructions prove particularly difficult to master in monolingual L1 acquisition. These structures, which stand on unsteady footing already in the native baseline, ought to be particularly vulnerable to reanalysis in heritage grammars. Thus, they are excellent candidates for the study of syntactic differences between monolingual and heritage speakers.

Bearing this vulnerability in mind, Pascual y Cabo (2013) targeted Spanish psych-verbs in a processing study that compared native and heritage, adult and child grammars. Cross-linguistically, psych-verbs denote a mental or emotional state, or the process that leads to such a state. These verbs are not uniform (e.g., Belletti and Rizzi, 1988; Landau, 2010); in Spanish, they fall into at least three classes. Pascual y Cabo concentrates on Spanish class III psych-verbs, among which *gustar* "like" is the most common. These psych-verbs are also referred to as reverse
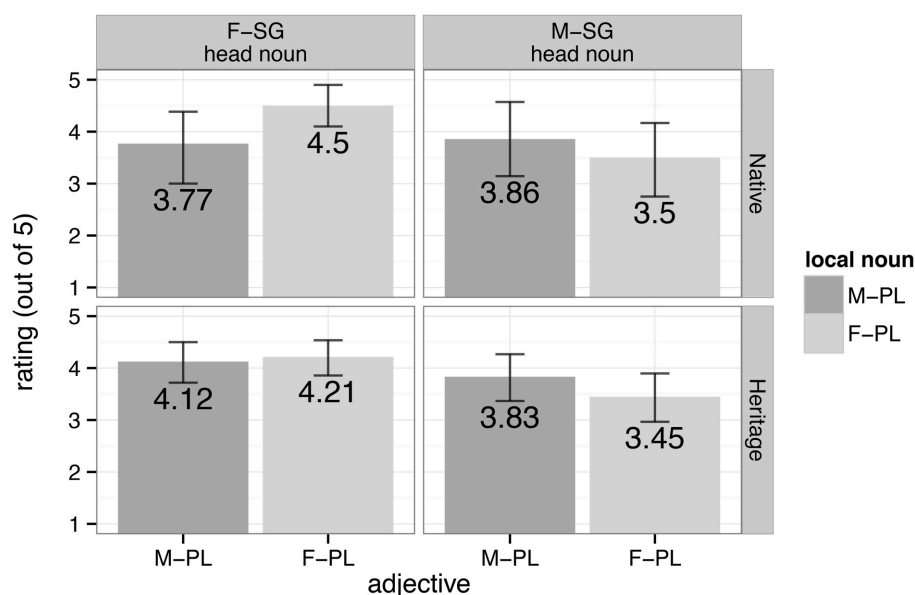
FIGURE 2 | Average ratings for potential attraction conditions from native (top) and heritage (bottom) speakers of Spanish in an extension of the methodology from Fuchs et al. (in press).

psychological predicates (RPP), owing to their non-standard argument mapping: the experiencer precedes the verb [*Katherine* in (5a)], but it is the post-verbal theme [*los kiwis* in (5a)] that is the syntactic subject of the sentence. Verbs of this type necessarily receive a stative reading. As strict statives, they expectedly resist passivization, as in (5b); syntactic accounts tie the lack of passivization to the absence of an agent-introducing *v*P projection in their argument structure (Belletti and Rizzi, 1988). Other classes of psych-verbs, namely those that allow agentive readings like *molestar* "bother" in (6a), can be passivized, (6b).

In a subsequent comprehension study, Gómez Soler (2012) determined that children as young as 3-years-old are able to comprehend this class of psych-verbs, but children's performance varied according to the specific verb used. Children performed remarkably well (at 79% accuracy) with *gustar*, but at chance (52%) with less common stative-only psych-verbs like *faltar* "lack." As is so often the case, different tasks yield different findings: a different comprehension study by Torrens et al. (2006) argued that children do not have adult-like understanding of these psych-verb constructions until around age 6;0. Although

(5)   *Spanish class III psych-verbs*
   a.   A   Katherine   le   gustan   los   kiwi-s.
        to   K.   3SG.DAT.CL   like.PRS.3PL   the.M.PL   kiwiPL
        'Katherine likes kiwis.'
   b.   *Los   kiwis   son   gustad-os   (por   Katherine).
        the.M.PL kiwiPL   be.PRS.3PL   like.PTCP-M.PL by   K.
        Intended: 'The kiwis are liked by Katherine.'

(6)   *Spanish class II psych-verbs*
   a.   Diana   molestó   a Adam.
        D.   bother.PST.3SG   to A.
        'Diana annoyed Adam (intentionally).'
   b.   Adam   es   molestad-o   (por   Diana).
        A.   be.PRS.3SG BOTHER.   PTCP-M.SG   by   D.
        'Adam is annoyed by Diana.'

This argument structure of stative psych-verbs has been the subject of much discussion in the literature on L1 and L2 acquisition of Spanish. Gómez Soler (2011) analyzes spontaneous child speech and shows that children start producing target-like *gustar* constructions quite early, at approximately age 1;10.

the exact time of acquisition of stative-only psych-verbs in Spanish is still up for debate, the evidence at hand supports the modest claim that they are acquired later by monolingual Spanish children than agentive predicates with regular argument-theta-role mappings.

Moving away from the native baseline, it should come as no surprise that these constructions also prove difficult for less idealized populations of learners. Regardless of the L1 of the speakers tested, psych-verbs with atypical argument structure consistently prove difficult for L2 learners of Spanish (Montrul, 1997; Quesada, 2008), although L2 learners eventually attain L1-level competency in producing and comprehending such constructions. With these facts in mind, Pascual y Cabo shifts attention to English-dominant heritage speakers of Spanish, who often lack formal schooling in their less dominant language. He notes that psych-verbs like *gustar* have two properties that make them vulnerable in the heritage grammar: their atypical argument structure, and the relative difficulty of their L1 acquisition. Based on a comprehension study of class III psych-verbs in Heritage Spanish, Pascual y Cabo hypothesizes that heritage speakers of Spanish reanalyze the psych-verb *gustar* to be optionally agentive, rather than strictly stative. In other words, heritage speakers might mistakenly align the argument structure of stative-only psych-verbs with less exotic agentive psych-verbs like *molestar.*

If this reanalysis were to take place, we should find evidence of it in passive constructions; this is precisely what Pascual y Cabo investigated. He predicted that if class III psych-verbs get reanalyzed as class II psych-verbs in heritage grammars, then heritage speakers would accept *gustar* and other such verbs in passive constructions. Native speakers, however, would find these constructions invariably unacceptable. The results of his acceptability judgment task confirmed this prediction: as expected, native speakers found passive constructions for stative-only psych-verbs to be categorically unacceptable, while heritage speakers at varying levels of proficiency rated these constructions as more acceptable. Pascual y Cabo argued that this result was sufficient to confirm his hypothesis that heritage speakers find *gustar* to be more compatible with passive constructions than native speakers do, and that this compatibility evidences the fact that heritage speakers are at least sometimes reanalyzing stative class III psych-verbs as agentive. Pascual y Cabo then considered the possible trajectory of this reanalysis. In order to determine whether the outcome implicated attrition, divergent attainment, or some other factor, Pascual y Cabo compared the performance of the original population of adult heritage speakers to child heritage speakers and child monolingual speakers, using the same acceptability task.

If the reanalysis of *gustar* were due to attrition, then at some earlier point in the lifespan of heritage speakers we would find more target-like behavior, which was lost on the way to adulthood (recall the discussion in Section *Developmental Trajectories of Heritage Speakers* above). Concretely, we would expect monolingual (and heritage) children to perform better at correctly judging passive *gustar* constructions to be unacceptable. However, this was not the case: both monolingual and heritage children performed worse than the adult heritage speakers. The fact that adult heritage speakers behave more like adult native speakers than do child monolingual speakers suggests that heritage speakers do improve their performance with these psych-verbs over time, and thus that the observed reanalysis does not arise from attrition. This improvement likewise suggests that divergent attainment is not the cause of reanalysis. Under

a divergent attainment story, we would expect similar behavior between child and adult heritage speakers.

Following Lightfoot (1991, 1999, 2012), Pascual y Cabo argues that "superficial performance innovations provided in the input from the immigrant generation contribute to the changes in H[eritage] S[peakers'] grammars" (Pascual y Cabo, 2013, p. 131). The original source, then, is attrition among L1 monolingual immigrants, who sometimes produce target-like *gustar* constructions, and sometimes do not. Next generation immigrant speakers (i.e., heritage language learners) receive this already non-standard input from their parents, which results in ambiguity in their mental representations of the syntax of the constructions at issue. The ambiguity forces heritage speakers to (economically) reanalyze the constructions, delivering the otherwise off-limits agentive constructions for psych-verbs.

The treatment of psych-verbs in heritage Spanish is clearly an innovation, the seeds of which are present in the native baseline, where verbs with non-canonical argument structure show a certain degree of instability. While it is clear that L1 speakers of Spanish ultimately acquire affective (experiencer) verbs, or at least *gustar,* the most prominent and frequent one among them, there are some Spanish dialects, for example in South America, where experiencers are expressed as subjects (not indirect objects; Anagnostopoulou, 1999); and there are other dialects where experiencers are encoded as direct objects (Franco, 1993, 1994). This variation indicates a certain degree of instability in the experiencer marking, exactly the instability that Pascual y Cabo picks up on in his description of the heritage speaker input. In addition, all heritage speakers of Spanish surveyed by Pascual y Cabo were dominant in English, which lacks similarly quirky subjects. Thus, even structural transfer from English may not be off the table as a possible contribution to reanalysis in these heritage speakers. Could we ever find instances of genuine reanalysis in adult heritage speakers, without transfer effects? We contend that such reanalysis is possible, and we turn to its example in the next section.

RELATIVIZATION: IN SUPPORT OF UNIVERSAL STRUCTURAL PRINCIPLES

Long-distance dependencies, relative clauses in particular, have long attracted the attention of linguists because they offer a window onto structural preferences in languages. If a language can relativize at a given position in the accessibility hierarchy in (7), then it can relativize at every position above it. To illustrate, if a language allows relativization of the oblique object, then we can expect the language to also allow relativization of the indirect object, direct object, and subject; if a language only allows one kind of relative clause, it will be a subject-extracted relative clause. Relative clauses also offer an excellent test case of memory constraints, which the parser needs to reckon with in the formation of long distance dependencies between the filler and its gap.

(7)   *Accessibility hierarchy* (Keenan and Comrie, 1977)
       subject > direct object > indirect object > oblique object > possessor > standard of comparison

Consider the subject-extracted relative clause in (8a), and the object-extracted relative clause in (8b). In both cases, the gap and the relative pronoun reference the subject of the matrix clause, *the reporter*.

(8) a.   The reporter$_i$ who$_i$ ___$_i$ harshly attacked the senator admitted the error.
     b.   The reporter$_i$ who$_i$ the senator harshly attacked ___$_i$ admitted the error.

Numerous studies have shown that, though (8a) and (8b) are grammatical and comprehensible, there are certain asymmetries regarding the ease (or lack thereof) with which speakers process these kinds of relative clauses. A large body of work continues to demonstrate that processing object-extracted relative clauses is more taxing, leading to increased processing times compared to subject-extracted relative clauses (see, for example, King and Just, 1991, for English; Frazier, 1987, for Dutch; Mecklinger et al., 1995, for Hungarian; Arnon, 2005, for Hebrew; Miyamoto and Nakamura, 2003, for Japanese; Kwon, 2008; Kwon et al., 2010, 2013; for Korean). Complementing the finding that object-extracted relative clauses are relatively costly to comprehend, recent work demonstrates that they are similarly costly to produce (Scontras et al., 2015).

Given the observed asymmetries in both production and comprehension costs, we might expect relative clauses to pose interesting issues for acquisition. (Recall from the previous case study the motivation for targeting psych-verbs as possible candidates for reanalysis: psych-verbs may be unstable in the native baseline, making them ideal candidates for reanalysis in

"A-movement" (i.e., movement to positions typically associated with arguments, like passivization), which seems to be the bane of developmental existence, and "A-bar movement" (i.e., the rest of movement, like relativization), which is acquired fairly unproblematically[4].

Assuming that relative clauses are more firmly established in the native baseline than psych-verbs, we might expect them to be less susceptible to change in heritage grammars. If relativization does not undergo the same processes of degradation that other areas of heritage grammars do—that is, if heritage speakers and native speakers perform equally well in comprehending and producing relative clauses—we would have support for the notion that competence in relativization is independent of quantity or quality of exposure. If, however, heritage speakers do diverge from native speakers in their performance with regard to relative clauses, then the observed differences may inform the trajectory of heritage grammars.

Polinsky (2011) used a picture-matching task to investigate the relativization behavior of English-dominant heritage speakers of Russian. English and Russian are both languages where native speakers can relativize at any point in the accessibility hierarchy [see the Russian examples in (9)]. The similarity between the two systems makes the examination of relative clauses in English-dominant heritage speakers of Russian particularly compelling, as it reduces the probability of transfer. However, unlike English, Russian has rampant scrambling (see King, 1995; Bailyn, 2004). Relative clauses are no exception: in both subject- and object-extracted relative clauses, the non-extracted noun phrase may occur either pre-verbally, (10a), or post-verbally, (10b)[5].

(9) a.   deti$_i$ [kotor-ye   ___$_i$ polučili   podarki   ot   babuški]
          children.NOM.PL   REL-NOM.PL   received   gifts.ACC.PL   from   grandma.GEN
          '(the) children that/who received gifts from Grandma'
     b.   podarki$_i$ [kotor-ye   deti   polučili   ___$_i$ ot   babuški]
          gifts.NOM.PL   REL-ACC.PL   children.NOM.PL   received   from   grandma.GEN
          '(the) gifts that the children received from Grandma'
     c.   babuška$_i$ [ot   kotor-oj   deti   polučili   podarki   ___$_i$]
          grandma.NOM   from   REL-GEN.SG   children.NOM.PL   received   gifts.ACC.PL
          'the grandmother from whom the children received gifts'

(10) a.   deti$_i$ [kotor-ye   ___$_i$ polučili   podarki]
          children.NOM.PL   REL-NOM.PL   received   gifts.ACC.PL
     b.   deti   [kotor-ye   ___$_i$ podarki   polučili]
          children.NOM.PL   REL-NOM.PL   gifts.ACC.PL   received
          '(the) children that received (the) gifts'

heritage grammars.) For relative clauses, however, the vast literature agrees that relative clauses do not pose any special difficulties in acquisition: Children acquire these constructions by the beginning of their third year (cf. Guasti and Cardinaletti, 2003, for Romance; Flynn and Lust, 1980; Hamburger and Crain, 1982; Diessel and Tomasello, 2000, for English; Friedmann and Novogrodsky, 2004, for Hebrew; Goodluck et al., 2006, for Irish; Slobin, 1986; Özge et al., 2009, 2010, for Turkish—the list goes on and on). The contrast between psych-verbs and relative clauses is part of a larger divide in the syntax literature between so-called

Given the similarities and differences between English and Russian, combined with the unique profile of abilities that

---

[4]More generally, the vulnerability of Spanish psych-verbs reflects difficulties in the acquisition of syntactic chains of arguments, in particular the acquisition of unaccusatives (e.g., Babyonyshev et al., 2001; Machida et al., 2004).

[5]The preverbal and postverbal positions in each type of relative clause are not totally equivalent, as they differ in terms of information structure; the right edge of the clause in Russian is strongly associated with focus (Adamec, 1966; Kovtunova, 1976; Padučeva, 1985). Studies of corpora find that these differences are reflected in the relative frequency of these types of RCs in Russian (Say, 2005; Polinsky, 2011; Levy et al., 2013).

characterizes heritage speakers, Polinsky's study was designed to answer two questions: first, does heritage Russian allow for the same expressivity in relativization structures, or have heritage speakers diverged from the native baseline in unnecessarily restricting themselves along the accessibility hierarchy? Second, does the presence of scrambling in the baseline Russian grammar (but not in the dominant English grammar) affect the grammar of relative clauses in the corresponding heritage language?

To answer these questions, Polinsky presented speakers with relativization structures that crossed two types of relative clause gaps (subject vs. object) with two orders of arguments in the relative clause (noun-verb vs. verb-noun). She predicted that subject-extracted relative clauses would be easier for heritage speakers to process than object-extracted structures, given the independently observed costs associated with object extraction; but she also expected the speakers would show effects of their dominant language. Specifically, Polinsky predicted that correspondences of surface order between certain Russian and English constructions would lead to differences between how heritage speakers and native speakers process scrambling within the relativization structures.

Participants were asked to choose between two pictures as they answered an auditory question with a relative clause in it. The stimuli all featured reversible actions, for example, chasing as in **Figure 3**. The question varied according to whether its relative clause featured subject vs. object extraction, and whether the order of arguments in the relative clause had been scrambled.
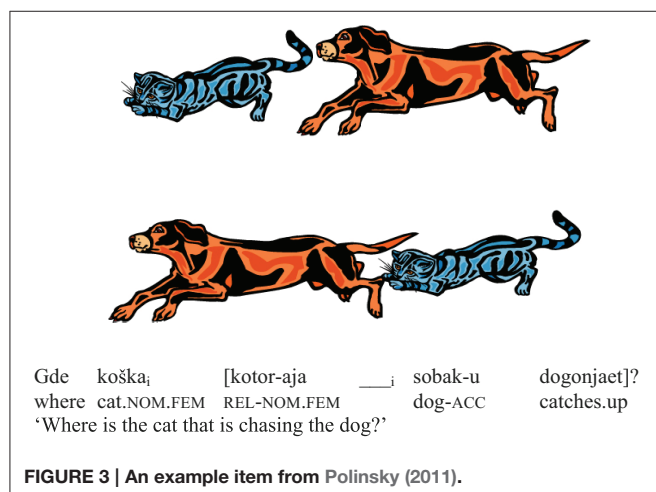
Polinsky's monolingual speakers, both adults ($n = 26$) and children ($n = 15$), found the task almost trivial, choosing the correct picture with ceiling-level accuracy. Heritage children ($n = 21$; average age 6;0) performed equally well. The surprising case was the performance of adult heritage speakers ($n = 29$), who exhibited a stark asymmetry in their performance between subject- and object-extracted relative clauses. These participants did perform quite well in subject-extracted identification tasks, but performed at chance when asked questions involving object extraction.

Polinsky argued for attrition as the source of the difference between native and heritage adult grammars. She noted that

both monolingual and heritage children performed essentially at ceiling, indicating that the adult heritage grammar could not be the result of a fossilized child language (i.e., divergent attainment), since the heritage children show perfect competence in this domain. Rather, these findings suggested that over their lifespan, the heritage speakers' competence with respect to relative clauses degraded, leaving the adult heritage speaker still capable of comprehending the easier subject-extracted relative clauses, but incapable of comprehending object-extracted relative clauses. Thus, Polinsky found evidence that relativization is not necessarily a robust area of linguistic competence: with reduced input and insufficient maintenance, competence in this area can become degraded. The observed attrition undoubtedly relates to a loss of morphological knowledge. If the heritage speakers did not process the nominative vs. accusative distinction, then they got no cue as to whether they were dealing with a subject- or object-extracted relative clause; they simply observed a clause with a transitive verb, a single overt argument, and a gap. In the absence of morphological cues, the default preference would then be to treat such a clause as a subject-extracted relative. However, this explanation alone cannot account for the comprehension of Russian relative clauses by heritage speakers, as there are also word order considerations to which we now turn.

It is natural to expect that the observed attrition may be caused by pressure from the dominant language, in this case English. If English were to blame, then relative clauses in which the internal word order mapped directly onto the word order of the analogous English sentence (i.e., relative clauses without scrambling) should have been easier for heritage speakers to process than ones in which the word orders did not match. The results of the study showed that this was not the case: heritage speakers performed equally well in identifying both subject-extracted configurations, and equally poorly in identifying both object-extracted configurations. Without any effect of scrambling on performance, we lack evidence of transfer from English. However, the absence of a scrambling effect suggests that heritage speakers were not entirely oblivious to the encoding of noun phrases, as morphology was the only cue to subject extraction in the scrambled relative clauses. Thus, Polinsky concluded that attrition in Russian heritage grammar, at least in the domain of relative clauses, is not the result of transfer. Instead, it is most likely the result of restructuring that occurs in the absence of sufficient maintenance. Ultimately, the heritage grammar is such that only subjects are accessible for relativization.

This evidence from Russian heritage grammars builds on and adds to several cross-linguistic discussions. The fact that heritage speakers performed uniformly well across subject-extracted conditions, and uniformly poorly across object-extracted conditions, regardless of word order within the relative clause, points to what has been labeled a "subject bias" observed in other syntactic environments (Keenan and Comrie, 1977; Kwon et al., 2010, 2013). Polinsky thus demonstrated that the privileged status of subjects amplifies in the heritage Russian grammar. The difference between native and heritage Russian speakers also conforms with the predictions of the accessibility hierarchy: native Russian speakers can relativize at all points on the hierarchy, whereas heritage Russian speakers can relativize

Gde	koška$_i$	[kotor-aja	___$_i$	sobak-u	dogonjaet]?
where	cat.NOM.FEM	REL-NOM.FEM		dog-ACC	catches.up

'Where is the cat that is chasing the dog?'

FIGURE 3 | An example item from Polinsky (2011).

at only one, the subject. This finding offers novel support to the reality of the subject as a linguistic category.

Like Pascual y Cabo (2013), this study also demonstrates the importance of comparing different age groups of heritage speakers in an effort to determine the trajectory of heritage grammars. Pascual y Cabo found that heritage adults performed better on the relevant task than children—evidence against attrition—whereas Polinsky made the same comparison but found, contrary to the expectations spelled out at the beginning of this section, that children performed better—evidence for attrition. This attrition is intriguing because it challenges the steady assumption that properties of movement (e.g., relativization), once acquired, should not be lost. It is clear, then, that a single result in one heritage group cannot be taken as evidence for a single process applying in heritage grammars across the board. Rather, in each grammatical domain and speaker population, a different combination of the factors is likely to be at play, shaping the heritage grammar.

AT THE INTERFACE: SCOPE INTERPRETATIONS

Even highly advanced multilingual speakers, be they L2 learners or heritage speakers, are known to demonstrate non-target-like linguistic behavior when they have to reason simultaneously about an internal component of the grammar and an external component (e.g., discourse; Sorace, 2011, and further references therein). This so-called "Interface Hypothesis" has been studied mostly in the domain of null subject licensing, where near-native speakers, heritage speakers included, perform less consistently[6]. In an attempt to expand the range of interface phenomena under consideration, our final case study reviews experimental findings on scope interpretations in heritage grammars.

Scope interpretations bring together at least three levels of representation: syntax (expressing the structural relationship among scope-bearing elements), semantics (expressing the logical implications of this structure), and pragmatics (supporting the expressed semantics and feeding back into the choice of syntax that determines it). We might therefore expect scope calculations to diverge from the native grammar in heritage speakers, as they perform the costly operation of integrating these various levels of linguistic representation. This divergence could take one of two paths: transfer from the dominant language resulting in an otherwise uncharacteristic pattern of behavior in the heritage speaker; or, faced with two systems of relatively different complexity, the simpler system winning out in the heritage grammar. Addressing these questions makes it necessary to test multiple systems; in addition to establishing baseline data in both languages, it is desirable to test heritage speakers' knowledge of scope in both the heritage language and their dominant language.

---

[6]The variation in near-native competency is determined by a number of factors, among which are the age of the onset of bilingualism (see Flores, 2010, 2012), the amount of input (see Montrul, 2016), and individual differences among speakers. In our discussion here, we abstract away from these additional factors.

Lee et al. (2011) take a step in this direction, trying to determine whether the grammar of scope in the heritage language could have an effect on the dominant language. The authors tested English-dominant heritage speakers of Korean on the interpretation of English negative sentences with universally quantified objects, as in (11). In English, this configuration yields ambiguity, corresponding to the scope of negation with respect to the universal quantifier.

(11) Mary didn't read all the books.
    a. *Surface scope* ($\neg > \forall$):
       It is not the case that Mary read all the books.
    b. *Inverse scope* ($\forall > \neg$):
       For each book, it is not the case that Mary read it.

Despite the availability of both surface and inverse interpretations for sentences like (11), speakers of English demonstrate a strong preference for surface interpretations. Presented with contexts supporting one or the other interpretation, native speakers of English accept inverse interpretations approximately 50% of the time (compared with a ceiling-level 90% acceptance rate for surface interpretations; Lee, 2009).

In Korean, similar sentences yield the opposite preference for interpretations (Han et al., 2007; O'Grady et al., 2009). Testing native speakers on sentences as in (12), Lee et al. (2011) show that surface interpretations yield near-50% acceptance rates, while inverse interpretations are accepted 90% of the time—the reverse of the English pattern.

(12) Mary-ka   motwun chayk-ul   anh ilk-ess-ta.
    Mary-NOM   all   book-ACC   not read-PST-DECL
    'Mary did not read all the books.'

Citing a processing explanation of these preferences from Grodner and Gibson (2005), Lee et al. suggest that differences in word order between English and Korean deliver the diverging patterns. In English, generating an inverse interpretation requires revising the initial parse, disrupting the linear operation of the processor and incurring a cost that results in a preference against the inverse, non-linear $\forall > \neg$ parse. Moreover, this inverse interpretation follows unambiguously from a ready alternative utterance: *Mary didn't read any books* (cf. the "pragmatic calculus" of Lidz and Musolino, 2006). In Korean, the SOV word order has this processor first encounter the universally quantified object, then negation; using the same reasoning used for English, we correctly predict the opposite preference, namely a preference for inverse interpretations in Korean.

Moving beyond the native baseline, Lee et al. tested the interpretation preferences of English-dominant heritage speakers of Korean *in English*. Their results show that these heritage speakers deploy their Korean preferences in English: 50% acceptance rate for surface vs. 90% for inverse. Perhaps surprisingly, early exposure to Korean seemed to interfere with scope calculation in the dominant language: English. Whatever its explanation, this result nevertheless raises important questions concerning the representation of scope in both monolingual and bilingual speakers. What aspect of the dominant English

grammar was affected by Korean? Unfortunately, Lee et al. did not test the scope preference of their heritage subjects in the heritage Korean grammar. Since that language was, at the time of the study, the weaker of the two in the subjects' bilingual repertoire, it is important to determine whether the scope preferences observed in monolingual Korean are still present in that language, when it is weakened by a dominant L2.

The study by Scontras et al. (2015) addresses these concerns by testing scope calculations by English-dominant heritage speakers of Mandarin in both of their languages, English and Mandarin. There is also another, more important difference between the two studies. Lee et al. demonstrate diverging preferences of scope interpretations between Korean and English in negative sentences with universally quantified objects. Crucially, speakers of each language allow both surface and inverse interpretations of these sentences, they merely prefer one interpretation over the other. However, assuming that Mandarin is a rigid surface scope language which completely disallows inverse scope in doubly-quantified sentences (an assumption which Scontras et al. test), comparing it with English, whose grammar permits inverse scope, allows for a fundamentally different comparison which more directly probes the robustness of each system as they intersect in the heritage grammar.

As in the previous case studies, the starting point is an establishment of the native speaker baseline. English sentences with more than one quantificational expression exhibit scope ambiguities. The ambiguities correspond to the relative scoping of the quantificational expressions at logical form. Various proposals deliver inverse scope; we focus on QR (May, 1977, 1985) for expository purposes and to align with discussions in previous experimental work on the topic. Under a QR approach, the surface and inverse interpretations of (13) follow from the schematic LFs in (13a) and (13b), respectively.

(13)   A shark attacked every pirate.
  a.   *Surface scope* ($\exists > \forall$):
       There was a single shark that attacked each pirate.
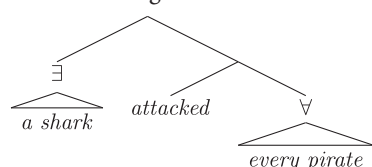
  b.   *Inverse scope* ($\forall > \exists$):
       For each pirate, there was a (different) shark that
                                                 attacked him.
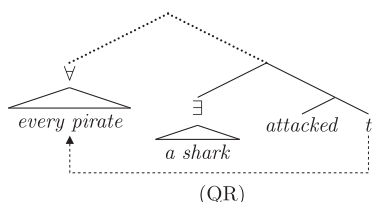
While speakers of English often accept inverse interpretations of doubly-quantified sentences, they display a reliable and robust preference for surface interpretations (cf. the preference for

surface scope in negative sentences; Tunstall, 1998; Anderson, 2004). This preference holds across a variety of dependent measures (e.g., measures of grammaticality like sentence ratings and truth judgments, or measures of processing difficulty), at a range of ages. Various proposals have been put forth to explain this preference, and they all share the feature that inverse scope calculation is costly relative to surface scope. The inverse LF in (13b) involves an additional step, covert QR of the object *every pirate* above the subject *a shark*. Because of this additional operation, the inverse LF, and thus the inverse interpretation, are more complex than the surface interpretation; because it is more complex, the inverse interpretation is the less preferred of the two.

Scontras et al. began by demonstrating these facts about scope preferences in native English, using a scene-description-naturalness rating task. Participants ($n = 114$) were asked to judge whether the sentence they heard appropriately described a co-occurring picture using a 7-point Likert scale (1 = "completely inappropriate," 7 = "completely appropriate"). The pictures matched either a surface (**Figure 4**, left) or an inverse (**Figure 4**, right) interpretation of the sentence[7]. **Figure 5** plots average ratings by condition; error bars represent bootstrapped 95% confidence intervals drawn from 10,000 samples of the data.

As expected, native English speakers allowed inverse scope in doubly-quantified sentences. However, these inverse interpretations came at a cost, resulting in lower ratings for inverse vs. surface interpretations. Still, the average rating of 4.46 (out of 7) for inverse scope was completely in line with preceding work on English scope; in general, complex structures are associated with lower ratings, and the ratings participants assigned in this task signal that inverse scope is not impossible, but simply less preferred.

In contrast to English, the picture in Mandarin Chinese appears remarkably stark. Since Huang (1982), many linguists have arrived at or accepted the conclusion that Mandarin does not allow inverse scope in doubly-quantified sentences. This prohibition means that Mandarin translations of the English sentences we considered reportedly allow only a surface interpretation. With respect to the scenarios depicted in **Figure 4**, (14) should therefore be judged true only with respect to the left image.

(14) You   yi-tiao   shayu   gongji-le   mei-yi-ge   haidao.
      exist one-CLF shark   attack-ASP every-one-CLF pirate
      'A/one shark attacked every pirate.'

Scontras et al. verified the claimed absence of inverse scope in Mandarin using the same sentence-picture naturalness rating task described above, this time testing native speakers of Mandarin ($n = 53$) on recorded sentences of Mandarin. **Figure 5** plots the results. Consistent with the received wisdom on inverse scope in Mandarin (*pace* Zhou and Gao, 2009), subjects demonstrated a strict resistance to inverse interpretations. Put simply, Mandarin does not allow inverse scope in

---

[7]The experimental pictures were taken from Benjamin Bruening's Scope Fieldwork Project: http://udel.edu/~bruening/scopeproject/scopeproject.html.
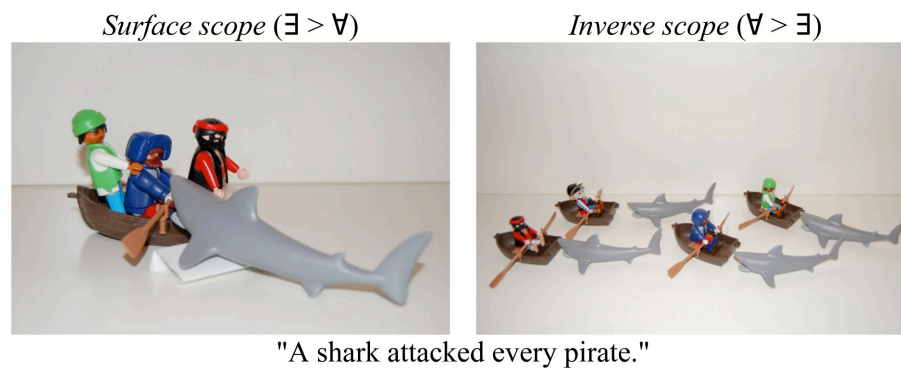
FIGURE 4 | An example item from Scontras et al. (in press).
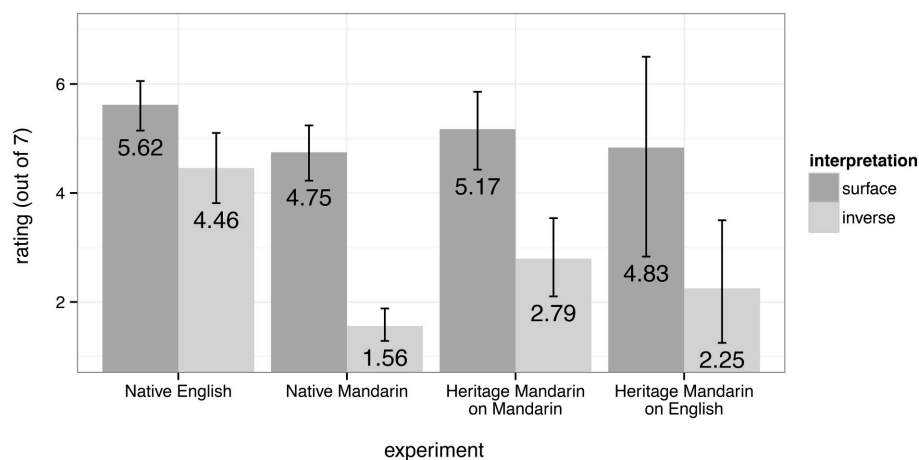
FIGURE 5 | Results from each of the four experiments from Scontras et al. (in press).

doubly-quantified sentences. This prohibition on inverse scope manifested as floor-level ratings, 1.56 out of a possible 7 points.

With clear baselines in hand—the availability of inverse scope in English and its absence in Mandarin—the authors then shifted their attention to the intersection of these two systems, namely English-dominant heritage speakers of Mandarin. What happens when one and the same individual presumably has access to both grammars?

Scontras et al. tested English-dominant heritage speakers of Mandarin on both the English ($n = 11$) and the Mandarin ($n = 26$) tasks described above, with the exception that the Mandarin task had instructions presented in English. The authors identified as heritage speakers those participants who learned Mandarin as their first language, but were dominant in English and lived in the United States at the time of testing. Results are plotted in **Figure 5** above.

Looking first at their scope in Mandarin, the picture that emerges suggests that these English-dominant heritage speakers of Mandarin did resist inverse interpretations for doubly-quantified sentences. Their ratings for the critical inverse condition were significantly lower than the English baseline for inverse scope (2.79 heritage Mandarin vs. 4.46 native English). However, heritage speakers' ratings were higher than the native Mandarin baseline (2.79 vs. 1.56 native Mandarin). One interpretation of these facts would have the heritage participants lacking inverse scope. The higher ratings for inverse conditions (relative to native speakers) would stem instead from the "yes-bias": heritage speakers are known to rate unacceptable/ungrammatical sequences higher than native controls (Benmamoun et al., 2013a,b; Laleko and Polinsky, 2013, in press)[8].

Another possibility is that the heritage speakers actually found inverse interpretations in Mandarin more acceptable than did native speakers, owing to transfer from their dominant language, English. We have seen that English allows inverse scope, so perhaps this possibility has permeated the heritage Mandarin grammar. The transfer of scope shifting would be incomplete,

---

[8]Second-language learners show a similar reluctance to reject clear grammatical violations. In their case, the lack of confidence can be attributed to their lack of implicit knowledge about many of the grammatical factors in play (Ellis, 2005, pp. 167–168).

owing to the lower ratings of inverse scope in heritage Mandarin compared to native English.

The final experiment from Scontras et al. proves crucial for teasing apart these competing hypotheses. Their results demonstrated that the English of these English-dominant heritage speakers of Mandarin does not allow inverse scope, or at least strongly resists it. These heritage speakers rated English inverse scope on average 2.25 out of a possible 7 points, a far cry from the 4.46/7 rating observed in the native English baseline. Given the observed lack of inverse scope in the English of English-dominant heritage speakers of Mandarin, it is unlikely that the intermediate ratings observed for heritage speakers tested in Mandarin stems from any transfer from a scope-allowing grammar. In fact, it would appear that these heritage speakers lack inverse scope in both their dominant English and their heritage Mandarin grammars.

By testing the robustness of the prohibition on inverse scope, the authors seem to have also tested the robustness of its permission: in the heritage speakers, even English lacked inverse scope. Could it be that the lack of inverse scope transfers from Mandarin to English in heritage speakers? Or might the relative expense of computing inverse scope, compounded with its reliance on a complex interaction between syntax, semantics, and pragmatics, render these interpretations too costly? We lack solid data to settle this question once and for all, but the authors present preliminary evidence from one last population which sheds some light on its answer: heritage speakers of English dominant in a language that prohibits inverse scope.

Given the global status of English and the prevalence of English-speaking communities, tracking down heritage speakers of English is not a trivial task. The target population for the present study is made more elusive by the requirement that these heritage speakers be dominant in a language that lacks inverse scope. Scontras et al. tested four Japanese-dominant heritage speakers of English living in Japan. Using the same English materials, these heritage speakers rated the critical inverse interpretations an average of 2.13 out of a possible 7 points. Taking into account the 4.46/7 baseline observed for native English, it appears that these heritage English speakers equally lack inverse scope. To summarize: of the four populations (native vs. heritage; English vs. Mandarin) and five grammars (native English, heritage English, native Mandarin, heritage Mandarin, and the English of heritage Mandarin speakers), Scontras et al. find just one clear case of inverse scope: the native English grammar.

Could it be that each of these heritage groups lose the ability for inverse scope because the rigid scope grammar is simpler? In fact, this is precisely what Lee et al. (2011) found for English-dominant speakers with early exposure to Korean. The confluence of evidence suggests that these bilinguals prefer simpler, less ambiguous grammars for scope—a preference visible in both the weaker and the dominant language. The authors fail to find interference from a dominant language when its system is more complex than the alternative. Instead, by expanding their sights beyond native grammars of scope, the authors found additional evidence for the precarious nature of scope calculations, manifested as a consistent pressure to simplify the grammar of scope: when two systems meet, the simpler system prevails.

If this simplification story is on the right track, the finding that heritage Mandarin speakers do not allow inverse scope in either of their languages does not necessarily entail that they have a robust Mandarin grammar. A grammar with ambiguity will be more complex than one without it: such ambiguities require abandoning a one-to-one mapping between surface structures and interpretations. The heritage Mandarin speakers that were tested might therefore have been more likely to adopt a Mandarin-like system, rather than the Mandarin system, because it is simpler, avoiding the added cost of inverse scope. In this sense, the change that resulted in the systems we observe was bidirectional, affecting both the English and the Mandarin systems. This resonates with observations, made mainly with respect to phonetics and phonology, according to which both languages in a bilingual system influence each other (cf. Flege, 1995; Flege et al., 1999, 2003; and see also Godson, 2003, for similar observations pertaining to heritage language). The results from scope thus offer novel support for the bidirectional interaction between two languages under contact.

CONCLUSIONS

The study of multilingualism has long been the intellectual property of linguistics subfields like sociolinguistics and language acquisition, and with good reason: we must understand the complexities of the multilingual experience before we can analyze its exponence in language users. With this limitation in mind, we began by considering the heterogeneity in just one sub-population of multilinguals, namely heritage speakers. With a clearer picture of the factors at play shaping the heritage grammar, we then presented case studies appropriating heritage language study into core domains of linguistic theory: the reorganization of morphosyntactic feature systems, the reanalysis of atypical argument structure, the attrition of the syntax of relativization, and the simplification of scope interpretations. In each case, we learned not just about the idiosyncrasies of the heritage grammar, but also about the native baseline and the resources and pressures at play in the development and maintenance of grammar.

We chose these case studies to highlight the breadth of heritage language research and its implications for linguistic theory, but we also chose them to evidence some useful methods in its practice. A few practical themes repeated themselves: establishment of a clear native baseline (a must for any comparison); determination of the input to heritage language acquisition by documenting the language of the parents (to locate the potential source of reanalysis and differences from the language in the homeland); determination of child heritage language behavior (to test for attrition over the lifespan); comparison of dominant and heritage language ability in the same population (to test for transfer, and its directionality). These practices help to narrow the possible explanations for observed atypical language behavior, pointing to both the trajectory and the outcome of grammatical phenomena in heritage speakers. And while these practices necessitate a good deal of time and care on

the part of the researcher, we have seen that they pay off, both by answering the specific questions targeted by the given study, and by raising additional questions central to any theory of grammar. We discuss two such questions in turn.

First, we have seen in most cases that the heritage grammar is often simpler than the native baseline with respect to the phenomenon of interest. But what does it mean to be simpler? This issue is related to two large and poorly defined notions in language science: complexity and default structures. These terms often arise in the context of sentence processing, where structures are shown to be more complex, or less default, on the basis of the processing profiles they elicit. But in the case of heritage linguistics, these terms take on a deeper meaning, one related to the grammar itself. Here we diagnosed complexity on a case-by-case basis, bringing to bear independent assumptions about language processing and architecture in the comparison of heritage and native grammars. If complexity is something that can be measured consistently, then we might expect heritage languages to consistently exhibit reduced complexity and thus reduced expressive power compared to the native baseline.

Which brings us to the second question, one we started this paper with: what does it mean to be a native speaker in the first place? Clearly the answer involves more than having L1-like phonology, which is typical of heritage speakers (Benmamoun et al., 2013b). But can we say more? On a practical note, answering this question, or at least recognizing it, is fundamental to researchers working on *understudied* and *endangered* languages. In many cases, such work involves bilingual consultants living in a dominant speech community other than the one of interest. The profile ought to ring familiar; these consultants stand a good chance of being heritage speakers of the language of interest. It is therefore possible, if not likely, that the language that gets documented will feature phenomena that are otherwise unexpected, and may seem challenging to universal principles of grammar. This issue was brought up, early on, in a seminal paper by H.-J. Sasse. He observed that differentiating native grammars "from the … situation of language decay is essential for the evaluation of data elicited from last generation speakers in a language death situation… How reliable is the speech of the last speakers [of a given community] and how much does it reveal of the original

structure?" (Sasse, 1992, p. 76). As we learn more about defining properties of heritage languages, this knowledge can be used to diagnose particular phenomena that indicate divergence from the baseline even in little-documented languages. Therefore, the significance of heritage languages lies not only in and of themselves. To illustrate, heritage languages are known to avoid embedded structures (Polinsky, 2008; Benmamoun et al., 2013b); the discovery of an exotic language without embeddings—the idealization of Pirahã, to some people—will be viewed to have completely different implications if this language is used just by a handful of remaining speakers, all of them heritage.

To conclude, we believe the value of the case studies we presented and many others that we lacked the space to mention serves as a signal that the need for myopathy in linguistic theorizing has left us. The time has come to embrace multilingualism; here we have proposed a specific way to do so: studying heritage languages. If nothing else, the reality that heritage speakers are everywhere multilingualism is cries out for a better understanding of their linguistic profile. More importantly, as we mentioned at the outset, the study of grammar is necessarily an indirect enterprise, achieved by studying the behavior of speakers. Why should we not help ourselves to as many speaker populations as possible, especially when a population presents novel data and new possibilities for asking and answering questions old and new? By approaching grammar from various entry points, we stand a better chance of moving our theories from the (specific) language-centric to the (general) Language-centric, the original aim of the Chomskyan enterprise.

ACKNOWLEDGMENTS

REFERENCES

Adamec, P. (1966). *Porjadok slov v Sovremennom Russkom Jazyke [Word Order in Modern Russian]*. Praha: Rozpravy Českosl. Akad. Ved.

Alcocer, P., and Phillips, C. (2009). "A cross-language reversal in illusory agreement licensing," in *Poster Presented at the 22nd Annual CUNY Conference on Human Sentence Processing* (Davis: University of California).

Alexiadou, A. (2004). "Inflection class, gender and DP-internal structure," in *Explorations in Nominal Inflection*, eds G. Müller, L. Gunkel, and G. Zifonun (Berlin: Mouton), 21–50.

Ammerlaan, T. (1996). *You Get it a Bit Wobbly…Exploring Bilingual Lexical Retrieval in the Context of First Language Attrition*. Unpublished doctoral dissertation, University of Nijmegen.

Anagnostopoulou, E. (1999). "On experiencers," in *Studies in Greek Syntax*, eds A. Alexiadou, G. Horrocks, and M. Stavrou (Dordrecht: Kluwer), 67–93.

Anderson, C. (2004). *The Structure and Real-time Comprehension of Quantifier Scope Ambiguity*. Unpublished doctoral dissertation, Northwestern University.

Anderson, R. (1999). Noun phrase gender agreement in language attrition. Preliminary results. *Biling. Res. J.* 23, 318–337. doi: 10.1080/15235882.1999.10162742

Antón-Méndez, I. (1999). *Gender and Number: Agreement Processing in Spanish*. Unpublished doctoral dissertation, University of Arizona.

Antón-Méndez, I., Nicol, J. L., and Garrett, M. F. (2002). The relation between gender and number agreement processing. *Syntax* 5, 1–25. doi: 10.1111/1467-9612.00045

Arnon, I. (2005). "Relative clause acquisition in Hebrew: toward a processing-oriented account," in *Proceedings of the 29th Boston University Conference on Language Development (BUCLD)*, eds A. Brugos, M. R. Clark-Cotton, and S. Ha (Somerville, MA: Cascadilla Press), 37–48.

Babyonyshev, M., Ganger, J., Pesetsky, D., and Wexler, K. (2001). The maturation of grammatical principles: evidence from Russian unaccusatives. *Linguist. Inq.* 32, 1–43. doi: 10.1162/002438901554577

Bailyn, J. F. (2004). Generalized inversion. *Nat. Lang. Linguist. Theory* 22, 1–49. doi: 10.1023/B:NALA.0000005556.40898.a5

Baker, M. C. (2008). *The Syntax of Agreement and Concord.* Cambridge: Cambridge University Press.

Behaghel, O. (1909). Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern [Relationships between size and ordering of constituents]. *Indogermanische Forsch.* 25, 110–142.

Belletti, A., and Rizzi, L. (1988). Psych-verbs and θ-theory. *Nat. Lang. Linguist. Theory* 6, 291–352. doi: 10.1007/BF00133902

Benmamoun, E., Montrul, S., and Polinsky, M. (2013a). Defining an "ideal" heritage speaker: theoretical and methodological challenges Reply to peer commentaries. *Theor. Linguist.* 39, 259–294. doi: 10.1515/tl-20 13-0018

Benmamoun, E., Montrul, S., and Polinsky, M. (2013b). Heritage languages and their speakers: opportunities and challenges for linguistics. *Theor. Linguist.* 39, 129–181. doi: 10.1515/tl-2013-0009

Bickerton, D. (1984). The language bioprogram hypothesis. *Behav. Brain Sci.* 7, 173–221. doi: 10.1017/S0140525X00044149

Bickerton, D. (1988). "Creole languages and the bioprogram," in *Linguistics: The Cambridge Survey*, ed F. J. Newmeyer (Cambridge: Cambridge University Press), 268–284.

Blake, R. (1983). Mood selection among Spanish-speaking children, ages 4 to 12. *Bilingual Rev.* 10, 21–32.

Bock, K., Carreiras, M., and Meseguer, E. (2012). Number meaning and number grammar in English and Spanish. *J. Mem. Lang.* 66, 17–37. doi: 10.1016/j.jml.2011.07.008

Bock, K., and Eberhard, K. M. (1993). Meaning, sound, and syntax in English number agreement. *Lang. Cogn. Process.* 8, 57–99. doi: 10.1080/01690969308406949

Bock, K., Eberhard, K. M., Cutting, J. C., Meyer, A. S., and Schriefers, H. (2001). Some attractions of verb agreement. *Cogn. Psychol.* 43, 83–128. doi: 10.1006/cogp.2001.0753

Bock, K., and Miller, C. A. (1991). Broken agreement. *Cogn. Psychol.* 23, 45–93. doi: 10.1016/0010-0285(91)90003-7

Bylund, E. (2009). Maturational constraints and first language attrition. *Lang. Learn.* 59, 687–715. doi: 10.1111/j.1467-9922.2009.00521.x

Carminati, M. N. (2005). Processing reflexes of the Feature Hierarchy (Person > Number > Gender) and implications for linguistic theory. *Lingua* 115, 259–285. doi: 10.1016/j.lingua.2003.10.006

Carstens, V. (2000). Concord in minimalist theory. *Linguist. Inq.* 31, 319–355. doi: 10.1162/002438900554370

Carstens, V. (2003). Rethinking complementizer agreement: agree with a case-checked goal. *Linguist. Inq.* 34, 393–412. doi: 10.1162/0024389033222 47533

Chomsky, N. (1965). *Aspects of the Theory of Syntax.* Cambridge, MA: MIT Press.

Cook, V. (2003). *The Effects of the Second Language on the First.* Clevedon: Multilingual Matters.

De Bot, K. (1990). "Language attrition, competence loss or performance loss," in *Sprache und Politik. Kongreßbeiträge zur 19. Jahrestagung der Gesellschaft für Angewandte Linguistik*, ed B. Spillner (Frankfurt; New York, NY: Peter Lang), 63–65.

Diessel, H., and Tomasello, M. (2000). The development of relative clauses in spontaneous child speech. *Cogn. Linguist.* 11, 131–151. doi: 10.1515/cogl.2001.006

Dubinina, I. (2012). *How to Ask for a Favor: An Exploration of Speech Act Pragmatics in Heritage Russian.* Unpublished doctoral dissertation, Bryn Mawr College.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language. *Stud. Second Lang. Acquis.* 27, 141–172. doi: 10.1017/S0272263105050096

Flege, J. E. (1995). "Second language speech learning: theory, findings, and problems," in *Speech Perception and Linguistic Experience*, ed W. Strange (Baltimore: York Press), 233–272.

Flege, J. E., MacKay, I. R. A., and Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *J. Acoust. Soc. Am.* 106, 2973–2987. doi: 10.1121/1.428116

Flege, J. E., Schirru, C., and MacKay, I. R. A. (2003). Interaction between the native and second language phonetic subsystems. *Speech Commun.* 40, 467–491. doi: 10.1016/S0167-6393(02)00128-0

Flores, C. (2010). The effect of age on language attrition: evidence from bilingual returnees. *Bilingualism Lang. Cogn.* 13, 533–546. doi: 10.1017/S136672890999054X

Flores, C. (2012). Differential effects of language attrition in the domains of verb placement and object expression. *Bilingualism Lang. Cogn.* 13, 533–546. doi: 10.1017/s1366728911000666

Flynn, S., and Lust, B. (1980). Acquisition of relative clauses: developmental changes in their heads. *Cornell Working Papers Linguist.* 1, 33–45.

Franco, J. (1993). Conditions on clitic doubling: the agreement hypothesis. *Añuario Del Seminario de Filología Vasca "Julio de Urquijo"* 27, 285–298.

Franco, J. (1994). The morphological, syntactic and semantic agreement status of Spanish object clitics. *Añuario Del Seminario de Filología Vasca "Julio de Urquijo"* 28, 707–789.

Frazier, L. (1987). syntactic processing: evidence from Dutch. *Nat. Lang. Linguist. Theory* 5, 519–559. doi: 10.1007/BF00138988

Friedmann, N., and Novogrodsky, R. (2004). The acquisition of relative clause comprehension in Hebrew: a study of SLI and normal development. *J. Child Lang.* 31, 661–682. doi: 10.1017/S0305000904006269

Fuchs, Z., Polinsky, M., and Scontras, G. (in press). The differential representation of number and gender in Spanish. *Linguist. Rev.*

Gass, S., and Selinker, L. (1992). *Language Transfer in Language Learning.* Amsterdam: John Benjamins.

Godson, L. (2003). *Phonetics of Language Attrition: Vowel Production and Articulatory Setting in the Speech of Western Armenian Heritage Speakers.* Ph.D. dissertation, University of California, San Diego.

Gómez Soler, I. (2011). "The L1 acquisition of Gustar: evidence against maturation," in *Selected Proceedings of the 4th Conference on Generative Approaches to Language Acquisition North America (GALANA 2010)*, ed M. Pirvulescu (Somerville, MA: Cascadilla Proceedings Project), 51–65.

Gómez Soler, I. (2012). "Acquiring Spanish psych-verbs: maturation or continuity?" in *Online Proceedings of the 36th Boston University Conference on Language Development (BUCLD 36)* (Somerville, MA: Cascadilla Proceedings Project).

Goodluck, H., Guilfoyle, E., and Harrington, S. (2006). Merge and binding in child relative clauses: the case of Irish. *J. Linguist.* 42, 629–661. doi: 10.1017/S002222670600421X

Grodner, D., and Gibson, E. (2005). Consequences of the serial nature of linguistic input for sentential complexity. *Cogn. Sci.* 29, 261–290. doi: 10.1207/s15516709cog0000_7

Guasti, M. T., and Cardinaletti, A. (2003). Relative clause formation in Romance child's production. *Probus* 15, 47–89. doi: 10.1515/prbs.2003.005

Haiman, J. (1983). Iconic and economic motivation. *Language* 59, 781–819. doi: 10.2307/413373

Hamburger, H., and Crain, S. (1982). "Relative acquisition," in *Language Development: Syntax and Semantics*, ed S. A. Kuczaj (Mahwah, NJ: Erlbaum), 245–274.

Han, C. H., Lidz, J., and Musolino, J. (2007). V-raising and grammar competition in Korean: evidence from negation and quantifier scope. *Linguist. Inq.* 38, 1–48. doi: 10.1162/ling.2007.38.1.1

Harley, H., and Ritter, E. (2002). Person and number in pronouns: a feature-geometric analysis. *Language* 78, 482–526. doi: 10.1353/lan.2002.0158

Harris, J. W. (1991). The exponence of gender in Spanish. *Linguist. Inq.* 22, 27–62.

Hartsuiker, R. J., Schriefers, H. J., Bock, K., and Kikstra, G. M. (2003). Morphophonological influences on the construction of subject-verb agreement. *Mem. Cognit.* 31, 1316–1362. doi: 10.3758/BF03195814

Huang, C.-T. J. (1982). *Logical Relations in Chinese and the Theory of Grammar.* Unpublished doctoral dissertation, Massachusetts Institute of Technology.

Hulsen, M. (2000). *Language Loss and Language Processing. Three Generations of Dutch Migrants in New Zealand.* Unpublished doctoral dissertation, University of Nijmegen.

Isurin, L., and Ivanova-Sullivan, T. (2008). Lost in between: the case of Russian heritage speakers. *Heri. Lang. J.* 6, 72–104.

Ivanova-Sullivan, T. (2014). *Theoretical and Experimental Aspects of Syntax-discourse Interface in Heritage Grammars.* Leiden: Bril.

Jarvis, S. (1998). *Conceptual Transfer in the Interlingual Lexicon*. Bloomington, IN: IULC Publications.

Keenan, E. L., and Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguist. Inq.* 8, 63–99.

Kim, J.-H. (2007). *Binding Interpretations in Korean Heritage Speakers and L2 Learners*. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.

King, J., and Just, M. (1991). Individual differences in syntactic processing: the role of working memory. *J. Mem. Lang.* 30, 580–602. doi: 10.1016/0749-596X(91)90027-H

King, T. H. (1995). *Configuring Topic and Focus in Russian*. Stanford, CA: CSLI Publications.

Kluender, R. (2004). "Are subject islands subject to a processing constraint?" in *Proceedings of the 23rd West Coast Conference on Formal Linguistics (WCCFL)*, eds V. Chand, A. Kelleher, J. Rodriguez, and B. Schmeiser (Somerville, MA: Cascadilla Press), 101–125.

Kovtunova, I. (1976). *Porjadok slov i Aktual'noe Èlenenie PredloŽenija.* [*Word Order and Information Structure of an Utterance*]. Moscow: Prosvešèenie.

Kramer, R. (2014). Gender in Amharic: a morphosyntactic approach to natural and grammatical gender. *Lang. Sci.* 43, 102–115. doi: 10.1016/j.langsci.2013.10.004

Kwon, N.-Y. (2008). *Processing of Syntactic and Anaphoric Gap-filler Dependencies in Korean: Evidence from Self-paced Reading Time, ERP and Eye-tracking Experiments*. San Diego, CA: University of California.

Kwon, N., Kluender, R., Kutas, M., and Polinsky, M. (2013). Subject/Object processing asymmetries in Korean relative clauses: evidence from ERP data. *Language* 89, 537–585. doi: 10.1353/lan.2013.0044

Kwon, N.-Y., Lee, Y., Gordon, P. C., Kluender, R., and Polinsky, M. (2010). Cognitive and linguistic factors affecting subject/object asymmetry: an eye-tracking study of pre-nominal relative clauses in Korean. *Language* 86, 546–582. doi: 10.1353/lan.2010.0006

Lago, S., Shalom, D. E., Sigman, M., Lau, E., and Phillips, C. (2015). Agreement attraction in Spanish comprehension. *J. Mem. Lang.* 82, 133–149. doi: 10.1016/j.jml.2015.02.002

Laleko, O., and Polinsky, M. (2013). Marking topic or marking case: a comparative investigation of heritage Japanese and heritage Korean. *Herit. Lang. J.* 10, 40–64.

Laleko, O., and Polinsky, M. (in press). Between syntax and discourse: Topic and case marking in heritage speakers and L2 learners of Japanese and Korean. *Linguist. Approaches Bilingualism*.

Landau, I. (2010). *The Locative Syntax of Experiencers*. Cambridge, MA: MIT Press.

Lee, M., Kwak, H.-Y., Lee, S., and O'Grady, W. (2011). "Processing, pragmatics, and scope in Korean and English," in *Proceedings of 19th Japanese-Korean Linguistic Conference*, eds H. Sohn, H. Cook, W. O'Grady, L. Serafim, and S. Cheon (Stanford, CA: CSLI Publications), 297–314.

Lee, S. (2009). *Interpreting Ambiguity in First and Second Language Processing: Universal Quantifiers and Negation*. Unpublished doctoral dissertation, University of Hawaii at Manoa.

Levine, G. (2001). *Incomplete First-language Acquisition in the Immigrant Situation: Yiddish in the United States*. Tübingen: Niemeyer.

Levy, R., Fedorenko, E., and Gibson, E. (2013). The syntactic complexity of Russian relative clauses. *J. Mem. Lang.* 69, 461–495. doi: 10.1016/j.jml.2012.10.005

Lidz, J., and Musolino, J. (2006). Why children aren't universally successful with quantification. *Linguistics* 44, 817–852. doi: 10.1515/LING.2006.026

Lightfoot, D. (1991). *How to set Parameters: Arguments from Language Change*. Cambridge, MA: MIT Press.

Lightfoot, D. (1999). *The Development of Language: Acquisition, Change and Evolution*. Oxford: Blackwell.

Lightfoot, D. (2012). Language acquisition and language change. *Wiley Interdiscip. Rev. Cogn. Sci.* 5, 677–684. doi: 10.1002/wcs.39

Lohndal, T. (2013). Generative grammar and language mixing. *Theor. Linguist.* 39, 215–224. doi: 10.1515/tl-2013-0013

Lorimor, H., Bock, K., Zalkind, E., Sheyman, A., and Beard, R. (2008). Agreement and attraction in Russian. *Lang. Cogn. Processes* 23, 769–799. doi: 10.1080/01690960701774182

Machida, N., Miyagawa, S., and Wexler, K. (2004). A-chain maturation reexamined: why Japanese children perform better on "full" unaccusatives than on passives. *MIT Working Papers Linguist.* 48, 91–112.

Martínez Mira, M. I. (2009). Spanish heritage speakers in the southwest: factors contributing to the maintenance of the subjunctive in concessive clauses. *Spanish Context* 6, 105–126. doi: 10.1075/sic.6.1.07mar

May, R. (1977). *The Grammar of Quantification*. Unpublished doctoral dissertation, Massachusetts Institute of Technology.

May, R. (1985). *Logical Form: Its Structure and Derivation*. Cambridge, MA: MIT Press.

Mecklinger, A., Schriefers, H., Steinhauser, K., and Friederici, A. (1995). Processing relative clauses varying on syntactic and semantic dimensions: an analysis with event-related brain potentials. *Mem. Cognit.* 23, 477–494. doi: 10.3758/BF03197249

Merino, B. (1983). Language loss in bilingual Chicano children. *J. Appl. Dev. Psychol.* 4, 277–294. doi: 10.1016/0193-3973(83)90023-0

Miyamoto, E., and Nakamura, M. (2003). "Subject/object asymmetries in processing relative clauses in Japanese," in *Proceedings of the 22nd West Coast Conference on Formal Linguistics (WCCFL)*, eds G. Garding and M. Tsujimura (Somerville, MA: Cascadilla Press), 342–355.

Montrul, S. (1997). On the parallels between diachronic change and interlanguage grammars: the L2 acquisition of the Spanish dative case system. *Spanish Appl. Linguist.* 1, 87–113.

Montrul, S. (2008). *Incomplete Acquisition in Bilingualism. Re-examining the Age Factor*. Amsterdam: John Benjamins.

Montrul, S. (2009). Knowledge of tense-aspect and mood in Spanish heritage speakers. *Int. J. Bilingualism* 13, 239–269. doi: 10.1177/1367006909339816

Montrul, S. (2011). First language retention and attrition in an adult Guatemalan adoptee. *Lang. Interact. Acquis.* 2, 276–311. doi: 10.1075/lia.2.2.05mon

Montrul, S. (2016). *Heritage Language Acquisition*. Cambridge: Cambridge University Press.

Montrul, S., and Sánchez-Walker, N. (2013). Differential object marking in child and adult Spanish heritage speakers. *Lang. Acquis.* 20, 109–132. doi: 10.1080/10489223.2013.766741

Odlin, T. (1989). *Language Transfer: Cross-linguistic Influence in Language Learning*. Cambridge: Cambridge University Press.

O'Grady, W., Lee, M., and Kwak, H.-Y. (2009). "Emergentism and second language acquisition," in *Handbook of Second Language Acquisition*, eds W. Ritchie and T. Bhatia (Bingley: Emerald Press), 68–81.

Otheguy, R., and Zentella, A. C. (2012). *Spanish in New York. Language Contact, Dialect Leveling and Structural Continuity*. Oxford: Oxford University Press.

Özge, D., Marinis, T., and Zeyrek, D. (2009). "Comprehension of subject and object relative clauses in monolingual Turkish children," in *Essays on Turkish Linguistics*, eds S. Ay, Ö. Aydin, I. Ergenç, S. Gökmen, S. Issever, and D. Peçenek (Wiesbaden: Harrassowitz), 341–350.

Özge, D., Marinis, T., and Zeyrek, D. (2010). "Production of relative clauses in monolingual Turkish children," in *Proceedings Supplement to the 34th Boston University Conference on Language Development (BUCLD 34)*, eds J. Chandlee, K. Franich, K. Iserman, and L. Keil (Boston, MA).

Padučeva, E. V. (1985). *Vyskazyvanie i Ego Sootnesennost' s Dejstvitel'nost'ju.* [*Utterance and its Reference*]. Moscow: Nauka.

Pallier, C. (2007). "Critical periods in language acquisition and language attrition," in *Language Attrition. Theoretical Perspectives*, eds B. Köpke, M. Schmid, M. Keijzer, and S. Dosterst (Amsterdam: John Benjamins), 99–120.

Pascual y Cabo, D. (2013). *Agreement Reflexes of Emerging Optionality in Heritage Speaker Spanish*. Unpublished doctoral dissertation, University of Florida.

Pascual y Cabo, D., and Rothman, J. (2012). The (Ill)logical problem of heritage speaker bilingualism and incomplete acquisition. *Appl. Linguist.* 33, 1–7. doi: 10.1093/applin/ams037

Pavlenko, A., and Jarvis, S. (2002). Bidirectional transfer. *Appl. Linguist.* 23, 190–214. doi: 10.1093/applin/23.2.190

Picallo, S. (1991). Nominals and nominalization in Catalan. *Probus* 3, 279–316. doi: 10.1515/prbs.1991.3.3.279

Pires, A., and Rothman, J. (2009). Disentangling contributing variables to incomplete acquisition competence outcomes: what differences across Brazilian and European Portuguese heritage speakers tell us. *Int. J. Bilingualism* 13, 211–238. doi: 10.1177/1367006909339806

Polinsky, M. (2006). Incomplete acquisition: American Russian. *J. Slavic Linguist.* 14, 191–262.

Polinsky, M. (2008). "Heritage language narratives," in *Heritage Language Education: A New Field Emerging* (New York, NY: Routledge), 149–164.

Polinsky, M. (2011). Reanalysis in adult heritage language: a case for attrition. *Stud. Second Lang. Acquis.* 33, 305–328. doi: 10.1017/S027226311000077X

Polinsky, M., and Kagan, O. (2007). Heritage languages: in the "wild" and in the classroom. *Lang. Linguist. Compass* 1, 368–395. doi: 10.1111/j.1749-818X.2007.00022.x

Potowski, K., Jegerski, J., and Morgan-Short, K. (2009). The effects of instruction on linguistic development in Spanish heritage language speakers. *Lanug. Learn.* 59, 537–579. doi: 10.1111/j.1467-9922.2009.00517.x

Quesada, M. L. (2008). "Yo gusto pasteles de chocolate: De la transitividad hacia la intransividad en la adquisición de verbos psicológicos en Español," in *Memoria del IX EILN*, ed R. M. Otiz-Ciscomain (Hermosillo, Mexico: Editorial Universidad de Sonora), 55–72.

Rakhilina, R., and Marushkina, A. (2014). "Non-calquing strategies in Heritage Russian," in *Paper Presented at the Second International Conference on Heritage/Community Languages* (Los Angeles, CA: University of California).

Ritter, E. (1993). Where's gender? *Linguist. Inq.* 24, 795–803.

Rothman, J., and Treffers-Daller, J. (2014). A prolegomenon to the construct of the native speaker: heritage speaker bilinguals are natives too! *Appl. Linguist.* 35, 93–98. doi: 10.1093/applin/amt049

Sasse, H.-J. (1992). "Language decay and contact-induced change: similarities and differences," in *Language Death: Factual and Theoretical Explorations with Special Reference to East Africa*, ed M. Brenzinger (Berlin: Mouton de Gruyter), 59–79.

Say, S. (2005). "Antipassives and A-relativization: between syntax and discourse pragmatics," in *Paper Presented at the 4th Typological School* (Tsakhkadzor).

Schmid, M. (2011). *Language Attrition*. Cambridge: Cambridge University Press.

Schwartz, B., and Sprouse, R. (1996). L2 cognitive states and the full transfer/full access hypothesis. *Second Lang. Res.* 12, 40–72. doi: 10.1177/026765839601200103

Scontras, G., Badecker, W., Shank, L., Lim, E., and Fedorenko, E. (2015). Syntactic complexity effects in sentence production. *Cogn. Sci.* 39, 559–583. doi: 10.1111/cogs.12168

Scontras, G., Polinsky, M., Tsai, C.-Y. E., and Mai, K. (in press). Cross-linguistic scope ambiguity: when two systems meet. *Lingua*.

Seliger, H. (1996). "Primary language attrition in the context of bilingualism," in *Handbook of Second Language Acquisition*, eds W. Ritchie and T. Bhatia (New York, NY: Academic Press), 605–625.

Silva-Corvalán, C. (1994). *Language Contact and Change: Spanish in Los Angeles*. Oxford: Oxford University Press.

Silva-Corvalán, C. (2003). "Linguistic consequences of reduced input in bilingual first language acquisition," in *Linguistic Theory and Language Development in Hispanic Languages*, eds S. Montrul and F. Ordóñez (Somerville, MA: Cascadilla Press), 375–397.

Silva-Corvalán, C. (2014). *Bilingual Language Acquisition: Spanish and English in the First Six Years*. Cambridge: Cambridge University Press.

Slobin, D. I. (1986). "The acquisition and use of relative clauses in Turkic and Indo-European languages," in *Studies in Turkish Linguistics*, eds D. I. Slobin and K. Zimmer (Amsterdam: John Benjamins), 277–298.

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. *Linguist. Approaches Bilingualism* 1, 1–33. doi: 10.1075/lab.1.1.01sor

Torrens, V., Escobar, L., and Wexler, K. (2006). "The acquisition of experiencers in Spanish L1 and the external argument requirement hypothesis," in *The Acquisition of Syntax in Romance Languages*, eds V. Torrens and L. Escobar (Philadelphia: John Benjamins), 181–202.

Tunstall, S. (1998). *The Interpretation of Quantifiers: Semantics and Processing*. Unpublished doctoral dissertation, University of Massachusetts, Amherst.

Valdés, D. (2000). "Introduction," in *Spanish for Native Speakers*, Vol. I. (New York, NY: Harcourt College), 1–32.

Vigliocco, G., Butterworth, B., and Garrett, M. F. (1996). Subject-verb agreement in Spanish and English: differences in the role of conceptual constraints. *Cognition* 61, 261–298. doi: 10.1016/S0010-0277(96)00713-5

Vigliocco, G., Butterworth, B., and Semenza, C. (1995). Constructing subject-verb agreement in speech: the role of semantic and morphological factors. *J. Mem. Lang.* 34, 186–215. doi: 10.1006/jmla.1995.1009

Vigliocco, G., and Franck, J. (1999). When sex and syntax go hand in hand: gender agreement and language production. *J. Mem. Lang.* 40, 455–478. doi: 10.1006/jmla.1998.2624

Vigliocco, G., and Nicol, J. L. (1998). Separating hierarchical relations and word order in language production: is proximity concord syntactic or linear? *Cognition* 68, B13–B29.

White, L. (1989). *Universal Grammar and Second Language Acquisition*. Amsterdam: John Benjamins.

Zhou, P., and Gao, L. (2009). Scope processing in Chinese. *J. Psycholinguist. Res.* 38, 11–24. doi: 10.1007/s10936-008-9079-x

Complexity Matters: On Gender Agreement in Heritage Scandinavian

Janne Bondi Johannessen[1]* and Ida Larsson[2]*

[1] MultiLing and Department of Linguistics and Nordic Studies, University of Oslo, Oslo, Norway, [2] Department of Linguistics and Nordic Studies, University of Oslo, Oslo, Norway

This paper investigates aspects of the noun phrase from a Scandinavian heritage language perspective, with an emphasis on noun phrase-internal gender agreement and noun declension. Our results are somewhat surprising compared with earlier research: We find that noun phrase-internal agreement for the most part is rather stable. To the extent that we find attrition, it affects agreement in the noun phrase, but not the declension of the noun. We discuss whether this means that gender is lost and has been reduced to a pure declension class, or whether gender is retained. We argue that gender is actually retained in these heritage speakers. One argument for this is that the speakers who lack agreement in complex noun phrases, have agreement intact in simpler phrases. We have thus found that the complexity of the noun phrase is crucial for some speakers. However, among the heritage speakers we also find considerable inter-individual variation, and different speakers can have partly different systems.

Keywords: Norwegian heritage language, Swedish heritage language, complexity, noun phrase, agreement, gender, declension class, attrition

INTRODUCTION

As has been shown in a number of recent studies (e.g., Montrul, 2008; Pascual y Cabo et al., 2012; Benmamoun et al., 2013; Kupisch et al., 2014; Johannessen and Salmons, 2015; Polinsky, 2015b; Larsson and Johannessen, 2015a,b), heritage languages can provide important insights into the nature of language acquisition, the linguistic effects of bilingualism across the lifespan of the speaker, and the principles behind linguistic change.[1] In earlier work (Larsson and Johannessen, 2015a,b; Larsson et al., 2015), we have identified four different factors that affect the development of the Scandinavian heritage language in America: contact between Scandinavian and English, contact between Scandinavian dialects (leading to dialect leveling and koineization), incomplete acquisition due to limited input and a language shift around the time of school start, and attrition. Attrition here refers to the loss of linguistic abilities that were once present in the speaker, due to lack of language use. Incomplete acquisition, on the other hand, refers to changes between generations, where the new generation acquires a grammar that is different in some respect from the grammar of the parents, due to limited or conflicting input (cf. Montrul, 2008, and see e.g., Sorace, 2004 who

---

[1] We use the term *heritage language* in the narrow sense to refer to a language acquired as a first language in a naturalistic setting, but in a society where it is not the dominant language. For heritage speakers, the first language will generally not continue to be the strongest, primary language after school start (cf. e.g., Polinsky, 2008 and references there). *Heritage Scandinavian* is here used to cover Heritage Norwegian and Heritage Swedish in America. For present-day Heritage Scandinavian speakers in America, English is always the primary language, and Scandinavian is weaker, secondary, and used only in a restricted set of situations and among a limited group of speakers (typically family members).

argues that it is important to distinguish what is lost in the language of the individual from features that were never there)[2].

These factors appear to play different roles in different linguistic domains. For instance, it has been shown that direct transfer from English affects the vocabulary (including function words), but not necessarily core syntax (Haugen, 1953; Hasselmo, 1974; Johannessen and Laake, 2012, forthcoming; Larsson et al., 2015). Larsson and Johannessen (2015a,b) argue that incomplete acquisition on the other hand has led to syntactic change: Heritage Scandinavian has a different word order in embedded clauses than do the Norwegian and Swedish varieties as spoken in Scandinavia. Attrition, we have argued, might, on the other hand, lead to loss of verb second in root clauses in some speakers (Eide and Hjelde, 2012; Johannessen, 2015a; Larsson and Johannessen, 2015b). The two syntactic changes thus seem to have different sources. In the former case, the adult heritage speakers pattern with pre-school L1 children; embedded word order is known to be difficult in L1 acquisition, too (see Larsson and Johannessen, 2015b and references there). In the latter case, the heritage speakers do not necessarily pattern with L1 learners, and the change is restricted to speakers that have not used their language regularly for many years, and who show other signs of attrition (most evidently, lexical retrieval delays). However, change that is due to attrition can be difficult to distinguish from change that should be understood in terms of incomplete acquisition, and it is likely that the two can be interrelated in individual speakers, and have similar results. There are differences, though. Attrition is expected to affect speakers that do not use their first language, regardless of the context of acquisition: Both heritage speakers and immigrant speakers can be affected equally. Among the Heritage Scandinavian speakers, we expect that incomplete acquisition affects most speakers in the community, since they typically have a very similar context of acquisition, but can have different patterns of language usage later in life. Incomplete acquisition is also expected to affect deeper grammatical properties in a different way than attrition, which most clearly affects processing and lexical retrieval (cf. Montrul, 2008).

In this paper, we look closer at one linguistic domain, the noun phrase. Scandinavian noun phrases clearly pose several difficulties in language acquisition, including double definiteness marking, agreement and gender assignment to the noun, and it has previously been shown to be affected in attrition. In a study of five young expatriate Swedes who had not spoken Swedish since childhood, Håkansson (1995) observed deviations in noun phrase–internal agreement in 35–68% (depending on the speaker) of the noun phrases, while word order (verb placement in main and embedded clauses) was target-like. In this respect, these heritage speakers behave differently from both L1 and L2 learners of Swedish; the latter typically show more deviations in word order than in morphology (see e.g., Pienemann and Håkansson, 1999). Other studies have confirmed that morphology is more sensitive to attrition than syntax. For

instance, it has been suggested specifically that gender in Heritage Norwegian is being attrited (Lohndal and Westergaard, 2014).

Here, we take a look at noun phrase agreement in Heritage Scandinavian speakers in America, with some comparison with immigrant speakers[3]. We focus on noun phrase internal agreement but also discuss noun declension. We will see that different groups of speakers produce deviations (relative to a baseline) in agreement, but few (if any) deviations with regard to declension class. We will also look at which forms are used, and what the reason for the deviations might be.

The paper is outlined as follows. Section Nominal Agreement and Declension Class in European Norwegian and Swedish gives an overview the relevant aspects of Scandinavian noun phrase morphosyntax and establishes a baseline. In Section Nominal Agreement and Declension Class in Heritage Norwegian and Heritage Swedish, we investigate American Norwegian and American Swedish respectively. In Section Results and Discussion, we discuss the patterns that we can observe with respect to differences between determiners and adjectives, the morphological forms that are used, and the role of complexity. We also briefly comment on the issue of attrition and acquisition. Section Conclusion gives the conclusion.

NOMINAL AGREEMENT AND DECLENSION CLASS IN EUROPEAN NORWEGIAN AND SWEDISH

In this section, we give a brief overview of noun phrase morphosyntax in European Norwegian and Swedish, which we assume, based on a cursory study of old recordings of American Norwegian and American Swedish, is the same in relevant respects to the language of the early immigrants (cf. Larsson and Johannessen, 2015b). This will form the baseline for our research on Heritage Scandinavian. When we discuss deviations in the language of the heritage speakers, these are understood in relation to the baseline. However, we maintain a rather liberal view of the baseline language, and only treat something as a deviation if the examples do not occur as dialect forms. In this way, we by necessity include in the baseline what was possibly present in the

---

[2]The term *incomplete acquisition* has caused some debate (see e.g., Pascual y Cabo and Rothman, 2012), and not everybody agrees that the two processes that the terms *incomplete acquisition* and *attrition* cover should be distinguished. See Larsson and Johannessen (2015b) for additional discussion.

[3]The data used in this paper is all from recordings done in the American Midwest in the 2010s. The recordings include naturalistic speech, with no elicitations, only recordings of sociolinguistic interviews between speakers and researchers, or conversations between speakers. Many of the Norwegian recordings have been transcribed and digitally processed, and are now available in a searchable corpus, Corpus of American Norwegian Speech (CANS), see Johannessen (2015b). At the time of the present study, the corpus contains the transcriptions, audio and video of the speech of 34 speakers, approximately 120,000 words. The Swedish recordings include both Heritage Scandinavian speaking descendants of the early immigrants, who are typically over 80 years old, and younger speakers that are descendants of more recent immigrants. A few speakers that emigrated themselves, and who we can refer to as immigrant speakers, were also recorded (see Andréasson et al., 2013 for an overview of the methodology); in the following we refer to their language as *Immigrant Swedish*. Only a few of the Swedish recordings have been transcribed, and they have not been included in the corpus. To establish the baseline, we have used older recordings of American Norwegian, collected by Einar Haugen in the 1940s–50s, and recordings of American Swedish collected by Folke Hedblom and Torsten Ordéus in the 1960s. For an overview of the older recordings, see Johannessen and Salmons (2012, p. 139) and Andréasson et al. (2013).

input of the first generation American-born heritage speakers, rather than what we know was in the input of our speakers[4].

There is considerable variation in nominal morphosyntax across the Scandinavian dialect continuum, for instance in the distribution of determiners, agreement on predicative adjectives, and in the gender system, some of which is relevant for the present study and will be discussed below. Importantly, it is possible to make generalizations that cover the Norwegian and Swedish varieties, and the variation that can be found is neither random nor unrestricted. Moreover, we have some knowledge of the dialect background of the heritage speakers in the study, and available dialect material (in particular the Nordic Dialect Corpus, Johannessen et al., 2009) makes it possible for us to check the American-Scandinavian data in relation to dialect speakers in Scandinavia. In the overview, we focus on the general features of Norwegian and Swedish nominal morphosyntax. (See Julien, 2005 for a thorough discussion of Scandinavian noun phrases, and Vangsnes et al., 2003 and Dahl, 2015 on dialect variation).

Determiners, Adjectives, and Agreement

Both Norwegian and Swedish have definite and indefinite articles (determiners), and there is a definiteness suffix (in addition to the prenominal definite article; see below on double definiteness). Determiners and adjectives are generally prenominal in both Norwegian and Swedish, see (1). Possessive pronouns are found both pre- and post-nominally depending on variety, but post-nominal possessives are infrequent in present-day Swedish. In Norwegian and Swedish, attributive adjectives, and determiners are inflected for number, definiteness, and, in indefinite singular noun phrases, for gender, as in (1). The system is represented by Norwegian in this section.

(1) a. *et   gammel-t hus*   (Norwegian)
         a.N.SG old.N.SG house

     b. *en   gammel-Ø hest*
         an.M.SG old.M.SG horse
         'an old horse'

     c. *ei   lit-a   elv*
         a.F.SG little.F.SG river
         'a small river'

In the present study, we focus on noun declension and noun phrase-internal gender agreement. Since some dialects lack predicative agreement (see e.g., Sandøy, 1988), predicatives are not considered here. In (2–4), we illustrate the agreement paradigms for determiners and adjectives. Note that in the Norwegian and Swedish adjectival paradigm, as in other Germanic languages, a distinction is made between so-called weak and strong inflection. Weak adjectival inflection is used in definite noun phrases, (2), and, in Mainland Scandinavian, does not distinguish gender and number. In this study, only

strong inflection of adjectives is included, in addition to singular determiners, which always show gender inflection.

(2) a. *den   gaml-e ku-a*   (Norwegian)
         the.F.SG. DEF old.DEF cow.F.SG.DEF
         'the old cow'

     b. *den   gaml-e hest-en*
         the.M.SG. DEF old.DEF horse.M.SG.DEF
         'the old horse'

     c. *det   gaml-e hus-et*
         the.N.SG.DEF old.DEF house.N.SG.DEF
         'the old house'

     d. *de   gaml-e hest-ene*
         the.PL.DEF old.PL horse.PL.DEF
         'the old horses'

Notice also that definiteness is marked in more than one place in all of the noun phrases in (2): both on the pre-posed definite article and on the nominal suffix. This way of marking definiteness is usually called double definiteness, and it distinguishes Norwegian and Swedish from e.g., Danish. Most definite noun phrases that have adjectival modification require double definiteness in Norwegian and Swedish, but the pre-posed article is otherwise not used in definite noun phrases; cf. (3) where the pre-posed article is not required (and, in fact, not possible unless the noun is modified by a relative clause or a preposition phrase).

(3) a. *ku-a*   (Norwegian)
         cow.F.SG.DEF
         'the cow'

     b. *hest-en*
         horse.M.SG.DEF
         'the horse'

     c. *hus-et*
         house.N.SG.DEF
         'the house'

Strong adjectival inflection appears in indefinite noun phrases. Gender is marked in the singular as in (1), but not in the plural (4).

(4) a. *gaml-e hest-er*   (Norwegian)
         old.PL horse.PL
         'old horses'

     b. *gaml-e hus-Ø*
         old.PL house.PL
         'old houses'

The Standard Swedish and Norwegian (Bokmål[5]) paradigms for the indefinite and the pre-posed definite article are given in **Table 1** below. The paradigm for adjectives is given in **Table 2**.

The only differences between different varieties of Swedish and Norwegian lie in the presence/absence of the feminine, and

---

[4]See Sorace (2004) for a comment on this potential methodological problem. In the present study, this is not a problem, we believe. Firstly, the deviations from the established baseline are not shared across the community. Secondly they show up even among present-day emigrant speakers, but they do not occur among L1 speakers in Norway or Sweden, nor in the older recordings of immigrant speakers.

[5]Norwegian has two official written standards, *Bokmål* and *Nynorsk*, and we have chosen to follow the former in the examples.

TABLE 1 | Definite and indefinite articles (determiners) in Norwegian (Bokmål) and Swedish.

		M.SG	F.SG	N.SG	PL
Indef.	No. *en*		*ei*	*et*	Ø
	Sw. *en* (common gender)			*ett*	Ø
Def.	*den*			*det*	*de*

TABLE 2 | Adjectival inflection in Norwegian and Swedish.

	M.SG	F.SG	N.SG	PL
Indef. ("strong" inflection)	No. Ø	-a*	-t	-e
	Sw. Ø (common gender)		-t	-a
Def. ("weak" inflection)	No. -e (for all genders and numbers)			
	Sw. -a (for all genders and numbers)			

*The Norwegian indefinite singular adjectival inflection is very rare, and only exists for a handful of lexemes.

in the fact that some varieties have the suffix *-e* (like No. Bokmål) while other varieties (like Standard Swedish) have *-a* on definite or plural adjectives[6].

Our cursory study of the older American Scandinavian recordings show that there is (as expected; cf. Section Gender below) some variation in the use of the feminine forms, and differences in the distribution of determiners (which are irrelevant for our purposes). Overall, the system is however identical to the system outlined above. We will therefore use this system as a baseline. If anything, the older heritage language has more morphological distinctions within the noun phrase than we have provided here, such as dative morphology (Johannessen and Laake, 2012), and not less. If we find examples of bare stems rather than inflected words, we therefore know that this is a deviation from the baseline.

Gender

In the following, we take gender to be an agreement category, distinct from declension class (following e.g., Corbett, 1991), but we make a distinction between gender agreement and gender assignment to the noun. In the latter, gender is generally assumed to be an inherent (lexical) property of nouns (e.g. Julien, 2005), but cf. e.g., Nygård and Åfarli (2013) who argue that gender is assigned to the noun in the syntax. Gender assignment is generally semantically opaque in Norwegian and Swedish, and it often has to be learned for individual lexical items (see e.g., Trosterud, 2001; Enger, 2014 for discussion).

The old Germanic three-gender system is retained to a lesser or higher degree. In Standard Swedish and some Norwegian varieties, the masculine and the feminine have collapsed into a common gender (see Fretheim, 1976/1985; Lødrup, 2011; Trudgill, 2013 for discussion). In many such varieties, feminine forms are retained in the nominal declension

only, with no feminine features on adjectives or determiners[7]. Instead, these nouns must be considered masculine/common gender, since their determiners and adjectival modifiers follow the masculine/common gender agreement pattern. In varieties with three-gender agreement, the indefinite determiner can be inflected in the feminine, as can a handful of adjectives (depending on variety), as exemplified in (1c). In varieties that only have two genders, the traditionally feminine nouns agree like (5), which is parallel to (1b):

(5)   *en   lit-en   elv*   (Norwegian)
     a.M.SG little.M.SG river
     'a little river'

Even in varieties of Norwegian and Swedish that retain the feminine, there can be a tendency toward a two-gender system, in which neuter remains as before, while masculine forms takes over at the expense of feminine forms.

If the Scandinavian Heritage language speakers have simplified their gender assignment system toward a default gender, it is likely that it will be toward the masculine. There are several reasons for this. First, the masculine/common gender is morpho-phonologically less marked than neuter. Both languages have Ø marking on the strong adjectival masculine inflection, while neuter singular has the suffix *–t*. Secondly, masculine nouns are more frequent than feminine nouns in the dialects that have the feminine (both with respect to type and token frequency; for Norwegian, see Heggstad, 1982, p. 12)[8]. In Swedish, common gender nouns are more frequent than neuter nouns (cf. Källström, 2008). More importantly, in the present context, masculine is the most popular gender for loanwords in the American Scandinavian varieties: for instance around 70% of the loan words in American Trønder[9] Norwegian are masculine (Hjelde, 1992, p. 84). Additionally, the masculine/common gender is also often overgeneralized in first and second language acquisition (see e.g., Rodina and Westergaard, 2013). Masculine has also been generalized in varieties that have changed from a three to a two-gender system. Moreover, in Swedish and Norwegian dialects that lack predicative agreement morphology in the plural, it is the masculine/common gender form that is used with plural predicative adjectives (Larsen and Stoltz, 1911, p. 45; Josefsson, 2009). Determiners like No. *sånn* 'such', which in some varieties do not have to agree when they have an abstract modal meaning, have the masculine agreement pattern as their default pattern (Johannessen, 2012)[10].

---

[6]Some varieties also have apocope, and drop the final vowel in some contexts (see e.g., Dahl, 2015, p. 135), but this is irrelevant for our study, which only looks at gender agreement marking, in which a final vowel exponent is never an option.

[7]The opposite is also true—there are Scandinavian varieties that show considerable reductions in the nominal declension system, collapsing the feminine and the masculine classes, but which still show evidence of a three-gender system in pronouns and possibly determiners. See Davidson (1990) for a discussion of the historical development in Swedish.

[8]Heggstad simply counted the frequency of each gender in a dictionary for Bokmål Norwegian, and found 54% masculine nouns, 21% feminine nouns and 25% neuter nouns.

[9]Trønder Norwegian is a cover term for the dialects spoken in the Trøndelag area and its surroundings in the middle part of Norway.

[10]Teleman (1969) argues against treating common gender as the default in Swedish, on the basis of the distribution of neuter singular morphology, which shows up in the absence of agreement in Swedish and Norwegian, as in other Indo-European languages. Teleman therefore concludes that it is the default. However,

Declension Class and Gender

In noun inflection, gender is visible on the form of the definiteness suffix (which has historically developed from a determiner) in the singular. In many varieties with a traditional three-gender system, nouns that have *-et* as a definiteness suffix are always neuter, while *-en* is masculine and *-a* feminine. In two-gendered Norwegian and Swedish, there are two general possibilities: In some varieties, *-et* is neuter, and *-en* and *-a* are common gender, whereas other varieties (such as Standard Swedish) has *-et* for neuter and *-en* for common gender. In language acquisition, the definiteness suffix in this way gives an unambiguous clue to the gender of the noun.

One complication is variation in the noun declensions of some old feminine nouns, as in (6). The noun in (6a) has a traditionally feminine form, but it cannot be considered feminine if it only triggers masculine/common gender agreement on determiners and adjectives. In such cases we will use the term declension class, and say that definite nouns ending in *-a* belong to a declension class that allows the *-a* suffix, and not to the feminine gender. We will take both forms to be target-like. (See Faarlund et al., 1997, p. 151 for more on the alternatives that exist.)

(6)   a. *elv-a*   (Norwegian)
            river.M.SG.DEF (still F.SG.DEF in many dialects)

        b. *elv-en*
            river.M.SG.DEF

In the plural, number inflection only partly reflects gender: Indefinite neuter nouns can have Ø-plural inflection in Swedish and Norwegian, and generally only old feminine nouns have–*or* as the plural suffix in Swedish (but not all do). That is, noun declension might give clues for gender in the acquisitional process, but plural suffixes do not necessarily reveal the gender of their hosts (see Källström, 1996; Enger, 2004 for discussion).

We will use the singular definiteness suffix as evidence for the gender of the noun in the heritage languages as long as we also find gender agreement morphology in the noun phrase. Thus, we follow Enger (2004) and assume that while there is in principle a distinction between gender (which necessarily affects associated words) and declension class, "it would be unwise to claim that the definite singular suffix is exclusively an exponent of declension and not at all of gender" (2004:65) in many varieties of Norwegian and Swedish, since this would leave several generalizations unexplained.[11] At the same time, an important part of the present study is to determine, for speakers whose agreement patterns are dwindling, whether target noun declension simply signals a declension class, or whether they indeed have gender distinctions, but that attrition has

caused deviant agreement patterns in their production (due possibly to processing problems related to lack of practice). The connection between gender and declension class will be discussed further in Section Gender, Agreement, and Declension Class.

NOMINAL AGREEMENT AND DECLENSION CLASS IN HERITAGE NORWEGIAN AND HERITAGE SWEDISH

In this study we focus on gender agreement. By *agreement* we mean cases where there is one or more exponents (possibly -Ø) of gender in the noun phrase apart from the noun (cf. e.g., Corbett, 1991; Baker, 2008). The examples we provided in Section Determiners, Adjectives, and Agreement above therefore illustrate gender agreement in Swedish and Norwegian[12]. We investigate attributive agreement inflection on determiners and adjectives as well as forms of the definiteness suffix in the singular. This means that we look at the presence or absence of suffixes that show agreement or non-agreement, as well as aspects of the nominal inflection. In noun phrases where more than one element would show gender agreement in the baseline (e.g., where both adjectives and determiners show agreement), all potentially agreeing forms have been counted separately, and these noun phrases are therefore represented more than once in the numbers.

If an adjective seemingly (given the baseline) does not agree with a suffixless noun, two analyses are possible: Either the adjective does not agree, or the noun has been assigned a different gender in that particular idiolect than in the baseline. The constructed examples in (7) illustrate this. (7a) will be interpreted as a deviation from the baseline with regard to agreement, since different choices have been made in the determiner and the adjective, while (7b) can also be interpreted as involving deviant gender assignment to the noun—unless we find evidence to the contrary—since there is agreement between the first two words, and the suffix-less noun does not reveal its gender. We treated these ambiguous cases separately at the onset. However, as we will see there is reason to assume that (7b), too, involves deviant agreement, given examples like (7c), where the noun is inflected, revealing its gender.

---

neuter forms do not seem to be the default spell-out of gender features, but rather, the default in the absence of gender features. We therefore maintain that masculine is the default value of gender features, and that -Ø is the elsewhere suffix for adjectival agreement.

[11] Again, this connection between gender and declension class does not hold for all varieties, and in some varieties the situation is changing, as discussed by Lødrup (2011) and Rodina and Westergaard (2015). As far as we know, the distinction between neuter and non-neuter is however maintained in the definiteness suffix.

[12] We do not include double definiteness marking in our study of agreement. One common analysis of double definiteness is to treat the prenominal determiner as a placeholder (e.g., Delsing, 1993), but it is also sometimes assumed that the two definite markers contribute partly different features (Julien, 2005). In other words, a noun phrase will here be considered target-like if double definiteness is missing, as long as the gender agreement is right. A few examples with modifiers in the neuter singular with plural nouns have been excluded from the study; neuter singular is sometimes used for mass nouns, or to mark absence of individuation (see e.g., Josefsson, Submitted). For instance, the Swedish speaker Annie produces examples like (i).

(i)   *mycket   finsk-or*
        much.N.SG Finnish.woman.PL
        'many Finnish women'

This is grammatical in many varieties of Swedish and Norwegian, given that the women are not individuated.

(7)   (Norwegian)

    a.   *en*   *god-t*
        a.M.SG good.N.SG
          gutt (target: *god* 'good'.M.SG.)
          boy. M.SG.
          (target: M.SG., but gender not visible on indefinite nouns)
        'a nice boy'

    b.   *et*   *god-t*
        a.N.SG good.N.SG
          gutt (target: *en*. 'a' M.SG, *god* 'good'.M.SG.)
          boy.M OR N?.SG
          (target: M.SG., but gender not visible on indefinite nouns. The speaker seems to have chosen N.SG in the agreeing words, thereby possibly having non-target gender assignment.)
        'a nice boy'

    c.   *det*   *god-e*
        the.N.SG.DEF good.DEF
          gutt-en (target: *den* 'the'.M.SG.DEF)
          boy.M.SG.DEF
          (target: M.SG., the target gender visible on the suffix of definite nouns)
        'the nice boy'

Closer investigation, considering inflected nouns in other sentences, reveals that the heritage speakers in fact do have the standard gender for those nouns. We therefore treat these as deviations in agreement morphology.

We have used partly different methods in the investigation of the Swedish and the Norwegian data, given the different types of data available. The Heritage Norwegian data are transcribed and available in a corpus (of 34 speakers), inviting more quantitative data (in addition to a closer study of two speakers). For Swedish, we have selected eight speakers with different backgrounds, but with Swedish as (one of) their L1[13]. We think that this combination of quantitative and qualitative data is possible since Swedish and Norwegian grammar share many of the relevant features. The different methods also contribute to the study in different ways, thus strengthening the results. The material is arguably rather small, but as we will see, we can still observe some patterns in the inter-individual variation, by looking in more detail at the individual speakers.

Agreement in the Corpus of American Norwegian Speech

The speakers in the Corpus of American Norwegian Speech (CANS, Johannessen, 2015b) are all born in the USA between the years 1900–1940. None of these speakers use their L1 very often, and they are all expected to show some signs of attrition (like lexical retrieval delays), in addition to changes due to

---

[13]We do not have full transcriptions for all of these speakers yet, so we also use the recordings. A few unclear examples have been disregarded.

incomplete acquisition or koinéization (cf. the introduction). (Due to the high number of speakers, we will not give more detailed information on these, except two of them in Section A Closer Look at Two of the Heritage Norwegian Speakers, unlike what we do for the Swedish heritage speakers in Section Heritage Swedish.)

We searched for the combination of (determiner)—adjective—noun in the corpus.

There are 171 hits of noun phrases where gender agreement is relevant. These are divided between 58 with the sequence adjective–noun (excluding those with a pre-adjectival determiner) and 113 cases for the sequence determiner–adjective–noun. There are altogether 21 cases that have non-target-like adjective or determiner agreement, but none that has a deviant definiteness suffix, see **Table 3** below.

There is a substantial difference in performance between the adjective-noun sequence and the determiner-adjective-noun sequence. (And notice that in the latter there can be non-target use of either of the two categories determiner and adjective, see Sections A Closer Look at Two of the Heritage Norwegian Speakers and Determiners and Adjectives) In the rest of the paper, we will refer to the combination of determiner-adjective-noun (sometimes with an omitted head noun, however) as a complex noun phrase, and to those with determiner-noun or adjective-noun as simple. The terms are here understood in a pre-theoretical sense—relating to the linear string, not hierarchical structure (see further Section Gender, Agreement, and Declension Class for discussion). As we shall see later, there are two different interpretations of the data in **Table 3**. Either noun phrases with determiners are more difficult, or it is complexity that matters. In the following sections we include determiner–noun sequences to investigate the two possibilities.

The complex noun phrases have 18% deviant constructions amongst all the speakers; a relatively high number. Some examples are provided below in (8):

(8)   (Heritage Norwegian)

    a.   *en*   *fin-t*
        a.M.SG.INDEF nice.N.SG.INDEF
          maskin (target: *fin* 'nice.M.SG.INDEF)
          machine.M.SG.INDEF
        'a nice machine' (Rushford_MN_01gm)

    b.   *ei*   *stor*
        a.F. SG.INDEF big.F/M. SG.INDEF
          famili (target: *en* 'a.M.SG.INDEF)
          family.M.SG.INDEF
        'a big family' (Harmony_MN_02gk)

    c.   *denna*   *andre*   *skolehuset*
        this.M.SG.DEF other.DEF school.building.N.SG.DEF
          (target: *detta* 'this.N.SG.DEF)
        'this other school building' (westby_WI_06gm)

In (8a), it is the gender on the adjective that deviates: the neuter form is used for the masculine (*fin*). In (8b) and (8c), it is the

TABLE 3 | Gender agreement and the definiteness suffix in Heritage Norwegian.

	Non-target agreement	Non-target definiteness suffix
Adjective-noun sequence	1/58 (2%)	0/58 (0%)
Determiner-adjective-noun seq.	20/113 (18%)	0/113 (0%)
Total	21/171 **(12%)**	0/171 **(0%)**

gender of the determiner that deviates from the baseline. In (8b) the feminine form of the indefinite article is used for the masculine (*en*). In (8c), the masculine form of the demonstrative (*denna*) is used for the expected neuter (*detta*)[14].

From the examples in (8) we see that the deviations involve several different forms, and that both adjectives and determiners can deviate. Deviant agreement can occur on determiner, adjective or both. However, there are some patterns. First, there are more deviations in determiners than in adjectives (cf. Section Determiners and Adjectives). Secondly, the neuter definite determiner *det* for the expected M/F determiner (or demonstrative) *den* occurs six times, whereas *den* is used twice for the neuter *det*. The feminine indefinite determiner *ei* is used twice for the expected masculine, and the masculine indefinite determiner is used three times in a neuter context. Thus, we cannot observe any clear generalization of the default in this data set, except that the non-target neuter definite singular determiner is used more often than the others. With respect to deviating adjectives, the picture is somewhat different: the masculine (corresponding to the bare stem) occurs for neuter six times, whereas neuter is used for M/F three times.

We have seen that Heritage Norwegian basically has target-like agreement, though there is 12% deviance in the relevant constructions. Looking closer at the speakers, however, it quickly becomes evident that there is considerable inter-individual variation. In fact, the majority of the speakers (20/34) in the CANS corpus only produce target-like examples, while 14/34 (41%) also produce non-targetlike examples. These speakers produce 86% of the hits, which means that while they make errors, they also produce a lot of utterances. There is also much individual variation in the frequency of non-targetlike NPs among the speakers that show deviations, ranging from 6% (3/52; coon_valley_WI_06gm) to 38% (5/13; Harmony_MN_02gk). This means that while many of the speakers in the corpus have baseline command of gender agreement, not all do. As a comparison, the adult Norwegian speakers (age 31–64) in Rodina and Westergaard (2015) show an accuracy of 99–100%.

[14]Following Johannessen (2008, p. 185–186), one could assume that there is no syntactic difference between the pre-posed article, the pre-posed possessive and the demonstrative in Norwegian; they are all determiners. Johannessen argues further that there seems to be individual variation between a system like the Norwegian one, and that of Danish, in which the demonstrative is not a determiner. This is contra Julien (2005) and Leu (2015), who argue that the proposed article, possessives and demonstratives have partly different syntax. Our data is too limited for us to investigate the behavior of the different elements in any detail, and for simplicity we refer to all three groups as determiners. For the present purposes, the particular analysis is not important, since it seems clear that what is relevant here is not the syntactic complexity of the noun phrase, but *linear* complexity (see further Section Gender, Agreement, and Declension Class).

It seems that for those speakers that produce non-target noun phrases, agreement poses problems, while gender assignment to nouns, as apparent from the definiteness suffix, is unproblematic. As noted above, examples like (9) could in principle be instances of an idiosyncratic gender assignment by the speakers, in which both *bilde* and *farmeår* are masculine rather than the target neuter. But the inflected forms of the noun *år* 'year' in the CANS corpus reveal that this word is rarely treated as masculine. The corpus gives 24 hits for the neuter *året* 'year.n.sg.def', and none for *åren* 'year.m.sg.def'. Searching for a simple noun phrase like *ett år* 'one year', we get 19 hits with a neuter determiner (*ett, æit, itt* etc.), and 8 with a masculine determiner (*ein, en, enn*). So both the definiteness suffix and the indefinite determiner reveal correct agreement forms compatible with the neuter-ness of the noun. Our hypothesis is that it is agreement rather than gender assignment and declension class that deviates for the heritage speakers. It is therefore worth asking if it is the complexity of the noun phrase that determines the extent of agreement.

(9)   a.   en god farmeår
         a.M good.M farming.year.N
           (target: *et* 'a'.N *godt* 'good'.N)
         'a good farming year' (decorah_IA_02gm)

   b.   en gammel bilde
         a. M old.M picture.N
           (target: *et* 'a'.N *gammelt* 'old'.N)
         'an old picture' (chicago_IL_01gk)

In the next section, we look closer at the two Norwegian Heritage speakers that have revealed the highest amount of deviant constructions, since these are the ones that are likely to have enough data to be subjected to a quantitative comparison of constructions and for possible systematicity to be evident.

A Closer Look at Two of the Heritage Norwegian Speakers

In the previous section we saw that amongst the 34 Norwegian heritage speakers in the CANS corpus, most have target-like agreement. Out of those that do not have full target-like agreement, there is great inter-speaker variation. In order to be able to investigate the linguistic properties that show some deviance, we need to look at speakers that produce some quantity of non-target agreement. There are two such speakers, Daisy (Chicago_IL-01gk) and Elsa (Harmony_MN_02gk). Daisy, 89 years old in 2010, was born in 1920 in Chicago by Norwegian immigrants, and Norwegian was spoken alongside English in her childhood home[15]. Her late husband did not speak Norwegian, and neither did her children. However, her father had lived with her until he died 15 years previously. She had not spoken Norwegian since. Daisy had been to Norway on five-six short trips. Elsa was born in 1930 in Spring Grove, all her grandparents had been born in Norway. She did not learn English until she started school. Her husband speaks Norwegian, but they never speak together. Two sons have settled in Norway with Norwegian

[15]When her maternal grandmother came from Norway to live with them, Daisy was 7. They spoke only Norwegian together. Daisy's Norwegian is definitely the dialect of her mother and maternal grandmother, originating in the town of Moss, Østfold.

TABLE 4 | Noun phrase morphology produced by the two most deviant Heritage Norwegian speakers.

	Deviant det-adj-noun agreement	Deviant det-noun agreement	Deviant noun-postposed poss. agreement	Deviant definiteness suffix
Daisy (Chicago_IL-01gk)	6/12 (50%)	4/51 (8%)	0/30 (0%)	0/59 (0%)
Elsa (Harmony_MN_02gk)	2/4 (50%)	2/23 (9%)	0/10 (0%)	0/39 (0%)
Total	8/16 **(50%)**	6/74 **(8%)**	0/40 **(0%)**	0/98 **(0%)**

families. She has been to Norway nine times, and has received several visits.

We can directly note that these two speakers have different gender systems. Whereas Elsa has a typical three-gender system, Daisy from Chicago has a Norwegian town dialect (reflecting, probably, her mother's dialect), and has only two genders: common gender (M/F) and neuter (N). Thus, we find both two-gender and three-gender systems among the speakers that show the most deviations. For these speakers, we have investigated the sequence determiner-adjective-noun, as well as the simpler (in a linear sense) phrases with determiner-noun and noun-possessive. We have also investigated all definite forms of their nouns. If these speakers reveal a difference between simple and more complex constructions, this will take us some way toward understanding the reason for the deviations. Notice that all pre-posed demonstratives, articles and possessives are regarded as determiners (see footnote 14) in Norwegian.

Table 4 sums up the findings for the two Norwegian speakers that have the most occurrences of deviant forms.

The table shows that there is indeed a difference between the more complex structures and the simplest ones, in which the most complex structures have 50% deviant agreement, while the much simpler determiner-noun structures have only 8% deviant forms, and the noun-possessive and noun-suffix structures have no deviance at all. Thus, it is quite clear that complex constructions are a challenge for these heritage speakers (though the absolute numbers are low). This result supports the trend in **Table 3**, Section Agreement in the Corpus of American Norwegian Speech, even if we have investigated only two speakers (though since the two speakers are also amongst the group of speakers in **Table 3**, we should not put too much emphasis on this).

In complex noun phrases, five of Daisy's non-target cases fall into the category of a non-target determiner, in which the neuter *det* has been chosen instead of the target masculine *den*. Elsa's non-target cases consist of one case similar to Daisy's, in which the neuter has been chosen for the masculine, and one in which a feminine indefinite article has been chosen instead of a masculine one, *ei* instead of *en*. There is only one non-target gender in the adjective inflection in these complex noun phrases, and there a neuter noun has been modified by a masculine determiner and a masculine adjective. We present complex noun phrases with non-target forms produced by our two informants in (10)[16].

(10c) was produced twice by Daisy, and once by Elsa. In (10d) there is both a deviant determiner and a deviant adjective.

(10) a. *det   siste   plass*
          the.N.SG.DEF last.DEF place.M.SG.INDEF
              (target: *den* 'the'.M.SG. DEF)
          'the last place' (chicago_IL_01gk)

      b. *det   eldste   John*
          the.N.SG.DEF oldest.DEF John. M
              (target: *den* 'the'.M.SG. DEF)
          'the oldest, John'[17] (chicago_IL_01gk)

      c. *det   første   gang*
          the.N.SG.DEF first.DEF time.M.SG.INDEF
              (target: *den* 'the'.M.SG. DEF)
          'the first time' (chicago_IL_01gk * 2, harmony_
              MN_02gk * 1)

      d. *en   gammel   bilde*
          a.M.SG.INDEF old. M.SG.INDEF picture.N.SG.INDEF
           'an old picture' (chicago_IL_01gk)
          (target: *et* 'a'.N.SG.INDEF, *gammelt* 'old'.N.SG. DEF)

      e. *ei   stor   familie*
          a.F.SG.INDEF big.M/F.SG.INDEF family.M.SG.INDEF
              (target: *en*.M.SG. INDEF)
           'a big family' (harmony_MN_02gk)

If we compare the complex noun phrase with a simple noun phrase (determiner–noun), we can see a clear difference, as evident from **Table 4**. In simple noun phrases, the two speakers achieve the target in the vast majority of the cases. There are 74 relevant cases, and of these 68 (92%) are target-like, i.e., only 8% are non-target-like, due to wrong gender agreement marking. This also shows that it is the complexity of the noun phrase that causes the difficulty with agreement, rather than e.g., difficulty with the forms of determiners *per se*. There are more deviant determiners in the complex noun phrases.

We can note that while we find an alternation between *den* 'the'.m.sg.def and *det* 'the'.n.sg.def in both types of noun phrases,

---

[16]Both speakers also produce target-like complex noun phrases:

(i)   den   store   bygning
      the.M.SG.DEF big.DEF building.M.SG
      'the big building' (chicago_IL_01gk)

(ii) den   eldste   jenta
      the.M/F.SG.DEF oldest.DEF girl.F.SG.DEF
      'the oldest girl' (harmony_MN_02gk)

[17]This example does not have the structure det+adj+noun, since the noun is appositive here. It is included, though, since there is a noun that determines the gender here, even though it is not expressed in the narrow noun phrase.

the number of non-target *det* is higher in the complex ones. We have, as shown in (12f), only one example of a non-target use of this determiner in simple noun phrases amongst the two speakers. But there are seven examples of the pre-posed, definite determiner *den* 'the'.m.sg.def that are target-like [see (11)], specifically five by Daisy and two by Elsa. Thus our hypothesis is strengthened; there are possible contexts (in simple noun phrases) where the non-target *det* 'the' n.sg.def could have appeared, but does not. (Thanks to a reviewer for this point).

Some target-like examples of the simple determiner-noun sequence are presented in (11). Notice that there is agreement in all three genders (two for Daisy), showing that gender, even for these speakers, is a category that is basically stable.

(11) a. *en*   *butikk* (Heritage Norwegian)
     a.M.SG.INDEF shop.M.SG.INDEF
     'a shop' (chicago_IL_01gk)

  b. *den*   *veien*
     that.M.SG.DEF way.M.SG.DEF
     'that way' (harmony_MN_02gk)

  c. *et*   *par*   *koner*
     a.N.SG.INDEF couple.N.SG.INDEF wives
     'a couple of wives' (chicago_IL_01gk)

  d. *dette*   *året*
     this.N.SG.DEF year.N.SG.DEF
     'this year' (harmony_MN_02gk)

  e. *ei*   *bok*
     a.F.SG.INDEF book.F.SG.INDEF
     'a book' (harmony_MN_02gk)

  f. *den*   *kirken*
     the.M.SG.DEF church.M.SG.DEF
     'that church' (chicago_IL_01gk)

Since, there are only six non-target determiner-noun sequences with these two speakers, we present all below in (12).

(12) a. *ei*   *brev*
     a.F.SG.INDEF letter.N.SG.INDEF
     (Heritage Norwegian) (target: *et*. N.SG. INDEF)
     'a letter' (harmony_MN_02gk)

  b. *ei*   *bryllup*
     a.F.SG.INDEF wedding.N.SG.INDEF
     (target: *et*. N.SG.INDEF)
     'a wedding' (harmony_MN_02gk)

  c. *en*   *fjell*
     a.M.SG.INDEF mountain.N.SG.INDEF
     (target: *et*. N.SG.INDEF)
     'a mountain' (chicago_IL_01gk)

  d. *en*   *barnebarn*
     a.M.SG. INDEF grandchild.N.SG.INDEF
     (target: *et*. N.SG.INDEF)
     'a grandchild' (chicago_IL_01gk)

  e. *det*   *slags*
     the.N.SG.DEF kind.M.SG.INDEF.GEN
     arbeid(target: *den*. M.SG.DEF)[18]
     work.N.SG.INDEF
     'the kind' (of work) (chicago_IL_01gk)

  f. *det*   *by*   (target: *den*.M.SG.DEF)
     the.N.SG.DEF town.M.SG.INDEF
     'the town' (chicago_IL_01gk)

Elsa from Harmony has two cases of non-target determiners; in both cases a feminine determiner has replaced the target neuter determiner.[19] The Chicago informant, Daisy, twice uses the masculine indefinite determiner to replace the neuter one. Conversely, she twice uses the definite neuter determiner *det* /de/ instead of the target masculine *den* /den/. However, she also has some examples of target use of the indefinite neuter determiner (*et par* 'a pair' *et hotell* 'a hotel'), and many examples of target definite masculine determiner, such as *den bygning* 'the building', *den dagen* 'the/that day'. There are several ways to interpret these results. One possibility is that Daisy finds agreement hard and mixes the forms. Another possibility is that she does not have the target gender assignment on the four relevant nouns. The latter can be checked. She uses the word *fjell* in the target definite form *fjellet* 'the mountain.' She also uses the correct plural definite determiner that characterizes neuter nouns: *barnebarna* 'the grandchildren.' Given, in addition to this, the large number of correct agreement forms, we take it that she does assign the correct gender to her nouns, and she knows basic agreement, but that agreement is sufficiently demanding for her to make non-target performance errors. However, when she chooses the form *det* /de/ instead of *den* /den/, it could have something to do with the phonological and semantic similarity to the English definite determiner *the*. We notice also that there is no hint of a resort to a default masculine gender. (See Section Determiners and Adjectives and Morphological Form and Type of Gender System for more discussion on these matters).

An equally simple pattern is the one with post-posed possessives, i.e., noun-determiner. Here we found 40 relevant hits (removing tagging errors and those where the determiner is invariable, like *hans* 'his' and *hennes* 'her'). In these cases, there

---

[18]One reviewer asks if maybe the determiner is targeting the noun *arbeid*, and asks what a monolingual corpus might show. We have searched for the non-target *det slags* and the target *den slags* in *Leksikografisk bokmålskorpus*, and found 1–one– example of *det slags*, but 1477 examples of *den slags*. It is clear that targeting a different noun here is not something that is done in monolingual language use.

[19]We choose to count that as non-target here, but this judgment is uncertain. Elsa's Norwegian ancestors are from the East of Norway (Østerdalen, Gudbrandsdalen, Ringerike, and Trøndelag), and in many of these dialects an unstressed neuter indefinite article is pronounced /ei/ or /i/, like the feminine article, instead of the stressed (and more standard) /et/. A relevant search in the Nordic Dialect Corpus (Johannessen et al., 2009) shows that this is the case in *inter alia* Kvam, Gausdal, Nordre Land, Brandbu, Åsnes, Drevsjø, Tolga, Røros, Dalsbygda, Gauldal, Meråker and Inderøy, which are all in the areas where she has ancestors. She never uses the article /et/. The speakers from the Norwegian areas mentioned all vary between a variant with –t, /et/ and one without, /ei/ or /i/.

was 100% correct score for the two informants. Examples are given in (13)[20], [21].

(13) a. *far*   *min*   (Heritage Norwegian)
      father.M.SG.INDEF my.M.SG
      'my father' (chicago_IL_01gk)

    b. *dattera*   *mi*
      daughter.f.sg.def my.F.SG
      'my daughter' (chicago_IL_01gk)

    c. *plassen*   *din*
      place.M.SG.DEF your.M.SG
      'your place' (chicago_IL_01gk)

    d. *mor*   *mi*
      mother.F.SG.INDEF my.F.SG
      'my mother' (harmony_MN_02gk)

Finally, the definiteness suffix is always target-like (see **Table 4**). Some examples are given in (14).

(14) a. *politiskolen*   (Heritage Norwegian)
      police.school.M.SG.DEF
      'the police school' (chicago_IL_01gk)

    b. *nabolaget*
      neighborhood.N.SG.DEF
      'the neighborhood' (chicago_IL_01gk)

    c. *hytta*
      hut.F.SG.DEF
      'the hut' (harmony_MN_02gk)

    d. *attenhundretalet*
      egihteenhundred.period.N.SG.DEF
      'the 19th century' (harmony_MN_02gk)

Heritage Swedish

The Norwegian data raises a number of questions that we will now address by looking in detail at the Heritage Swedish data. Firstly, some of the inter-speaker variation might be due to what types of utterances the speakers produce. That is, Elsa and Daisy might show more deviations because they produce a higher number of complex noun phrases than other speakers. By considering all utterances produced by a number of Heritage Swedish speakers, we can address this question. Secondly, the corpus data gives us few direct clues to what the reason is that some speakers deviate to a higher or lesser extent. (But we know from a previous study (Johannessen, 2015a) that Daisy shows

---

[20]From (13b) it looks as if Daisy has retained the feminine gender after all. As pointed out by Fretheim (1976/1985) and Lødrup (2011) post-nominal possessive forms are special since they can retain the old feminine morphology although the three-gender system has otherwise disappeared. Lødrup shows that these forms should not be treated as feminine agreement forms in modern varieties of Oslo Norwegian. Whether his analysis can be extended to speakers like Daisy, or if the form *mi* shows that Daisy has some remnants of the three-gender system is not clear from the available data.

[21]The postposed possessives in (13) show that the heritage speakers have a good command of another part of the grammar, too. Examples (13a,d) show that they are aware that certain kinship terms (especially *mor* 'mother', *far* 'father' and *bror* 'brother") can occur with postposed possessives with an indefinite form of the noun, unlike other nouns, as in (13b,c).

TABLE 5 | The American-Swedish speakers.

Name	Born	Nationality	Speaking pattern
Annie	1931 in Northern Sweden (Norrbotten)	Emigrated in 1949	Married to Norman, speaks Sw. in some contexts
Martin	1958 in Northern Sweden (Dalarna)	Emigrated in 1968	Speaks Sw. very rarely
Arthur	1929 in Minnesota	Sw. monoling. until school start, much Swedish at home	Grew up with Swedish grandparents, speaks Sw. on occasion
Albert	1921 in Minnesota	Sw. as L1, En. at age ten	Speaks Sw. on occasion
Norman	1930 in Chicago	Sw. monoling. until the age of 5	Married to Annie, hears Sw., speaks rarely
Konrad	1933 in Minnesota	Sw. monoling. until school start	Speaks Sw. very rarely
Theodor	1922	Sw. and En. as L1	Speaks Sw. on occasion
Amos	1921 in Minnesota	Sw. and En. as L1	Has not spoken Swedish since 1976

other signs of attrition.) In the Swedish data, we have access to different groups of speakers—both American-born heritage speakers and immigrant speakers—and by considering their linguistic background, we have a further way of addressing this question, too. Thirdly, by considering more data, and by looking at the inter-speaker patterns, we might find clearer patterns with respect to which forms are overgeneralized and when the deviations occur. We will see that the data from Swedish also supports the conclusion above that deviations involve non-target agreement, rather than non-target gender assignment. Recall that we expect many of the Heritage Swedish speakers to have a two-gender system where the singular definiteness suffix without exception signals the gender of the noun.

Since the background of the speakers is of some importance, we start by giving an overview of this, before we turn to agreement patterns and declension.

The Heritage Swedish Speakers

For Swedish we have investigated all noun phrases produced by eight speakers with partly different backgrounds. Two of the speakers (Annie and Martin) are first generation immigrants. In other words, they speak what we could refer to as Immigrant American Swedish. We can assume that they acquired Swedish fully in their childhood. Deviations should therefore be due to attrition rather than incomplete acquisition. The other six speakers are born in the American Midwest. Four of them (Arthur, Albert, Norman, Konrad) were monolingual in Swedish until around the age of 5–6. Of these, Arthur and Albert continued to have Swedish as the dominant language the longest. Albert reports that he grew up with his grandparents who spoke little or no English, and in the fifth grade his teacher told him to start speaking English. Two speakers (Amos and Theodor) report that they were early bilinguals, with both English and Swedish before school start. Annie and Norman are married, and they were interviewed together. A summary is given in **Table 5** below.

Like most of the speakers in the American Norwegian corpus, none of the speakers use their Swedish regularly. Amos last spoke Swedish when he visited Sweden in 1976 (36 years ago at the time of the interview). He writes a little Swedish, and we have received some e-mail communication from him. During the interview, Norman is reluctant to speak Swedish, but he has less difficulty after a few minutes. His wife Annie speaks Swedish to some of her friends, and Norman seems more used to listening to Swedish than to speaking it.

Of the two immigrant speakers, Martin has an earlier and more abrupt onset of English than Annie. He emigrated at the age of 10 and reports that he learnt English immediately and without difficulty. He also reports that he no longer speaks Swedish with his parents. At the same time, he behaves like a native speaker in a different way than the American-born heritage speakers. For instance, his judgment of embedded word order is stable and native-like, and on occasion, he uses Swedish derivational morphology productively and correctly to compensate for his lack of vocabulary.

Noun Declension

We saw above that the Norwegian heritage speakers never have deviations in the form of the singular definiteness suffix. The same is true for the Swedish speakers: Only one of the 201 examples in the data shows a deviation. Some of the target-like examples are given in (15a–d), and the deviating example is given in (15e)[22]. The data is summarized in **Table 6**.

(15)   a.   *brev-et* (Amos) (Heritage Swedish)
           letter.N.SG.DEF

      b.   *golv-et* (Arthur)
           floor.N.SG.DEF

      c.   *konditori-et* (Norman)
           cafe.N.SG.DEF

      d.   *student-en* (Norman)
           student.C.SG.DEF

      e.   *krig-en* (Norman)
           war.C.SG.DEF (target: *krig-et* 'war.N.SG.DEF')

However, from this data alone we do not know if the gender system is intact in Heritage Swedish. While the singular definiteness suffix unambiguously signals the gender of the noun in Standard Swedish (also reflected in agreement), it might mark only declension class for these heritage speakers. The Norwegian data suggested that the problem is really agreement, not gender assignment. In the next section, we will see that the same holds for Swedish.

Gender Agreement in Complex and Simple Noun Phrases

We have investigated all noun phrases where gender agreement is expected; the results are summarized in **Table 7**. As we

---

[22]Some speakers have the definiteness suffix *-a*, in addition to *-en*. For instance, Konrad uses both the form *boken* and the form *boka* for 'the book'. This is not unexpected, since it is common in many Swedish dialects (as it is in Norwegian). The distribution is targetlike. Both *-a* and *-en* are restricted to common gender nouns. There is no evidence for feminine agreement in the language of Konrad.

TABLE 6 | The definite singular form of nouns in Heritage Swedish.

Name	Non-target definiteness suffix
Annie	0/22 (0%)
Martin	0/24 (0%)
Arthur	0/13 (0%)
Albert	0/25 (0%)
Norman	1/17 (6%)
Konrad	0/64 (0%)
Theodor	0/13 (0%)
Amos	0/23 (0%)
Total	1/201 **(0.5%)**

did for the Norwegian data, we distinguish between complex noun phrases (with determiner-adjective-noun) and simple noun phrases (adjective-noun or more often determiner-noun). The simple noun phrases include two examples with a post-nominal possessive (both produced by Arthur), which is possible in a few Swedish dialects. In principle, many of the deviations in simple noun phrases (and some of the complex ones) are ambiguous between deviating gender agreement and gender assignment. As we have argued for Norwegian, these cases should most likely generally be treated as involving deviating agreement. This will be discussed further below.

The overall frequency of non-target agreement is 10%, and thus very similar to what we found for Norwegian (12%). As for Norwegian, we find a difference between complex and simple noun phrases: the former show deviations in 16% of the cases, whereas the latter have 7% deviating forms. However, complexity is not a clear factor for all speakers; One speaker (Arthur) has the opposite pattern. Two speakers (Albert and Konrad) are target-like (with a single exception) in both types of noun phrases.

The results reveal considerable inter-individual variation also in other respects. Some speakers show no or almost no examples of deviations, whereas others have a considerable amount. Again, this is what we saw for Norwegian. The variation can give us some insight into the factors that cause the deviant forms. We can note that the two immigrant speakers show partly different behavior. Annie produces only two deviant forms. They occur in the same noun phrase, and this particular example is clearly due to lack of planning: She immediately afterwards switches from the neuter noun *ställe* 'place' to non-neuter *by* 'village'. Martin, on the other hand, shows a few examples of deviations that suggest some occasional difficulty with nominal agreement. In all of these cases, the default (C.SG) is generalized on determiners or adjectives, as in (16):

(16)   (Immigrant Swedish)

      a.   *en*   *annat*
           an.C.SG.INDEF other.N.SG.INDEF
              stålverk (target: *ett* 'an'.N.SG.INDEF)
              steelworks.N.SG.INDEF
           'a different steelworks' (Martin)

TABLE 7 | Gender agreement in noun phrases in Heritage Swedish.

Name	Non-target agreement in complex noun phrases Det–A–N	Non-target agreement in simple noun phrases Det–N or A–N	Non-target agreement in total
Annie	2/45 (4%)	0/24 (0%)	2/69 (3%)
Martin	5/22 (23%)	1/44 (2%)	6/66 (9%)
Arthur	0/13 (0%)	2/17 (12%)	2/30 (7%)
Albert	0/2 (0%)	0/13 (0%)	0/15 (0%)
Norman	6/9 (29%)	1/16 (17%)	7/25 (28%)
Konrad	0/13(0%)	1/81 (1%)	1/94 (1%)
Theodor	4/15 (27%)	3/21 (14%)	7/36 (19%)
Amos	4/11 (36%)	8/24 (33%)	12/35 (34%)
Total	21/130 **(16%)**	16/240 **(7%)**	37/370 **(10%)**

b. *en   gammal*
   an.C.SG.INDEF old.C.SG.INDEF
       stålbruk
       steelworks.N.SG.INDEF
   'an old steelworks' (Martin)
   (target: *ett*   an.N.SG.INDEF,   *gammalt*   old.
       N.SG.INDEF)

In principle, (16b) could be interpreted as a deviation in gender assignment rather than agreement, as we have noted. However, Martin otherwise has no deviations in noun inflection, and he also produces the target-like neuter definite inflection of the noun *stålbruk* 'steelworks' (17). As we have seen, this is the general pattern among the speakers.

(17)   *stålbruk-et*   (Immigrant Swedish)
       steelworks.N.SG.DEF (Martin)

Both Annie and Martin can be assumed to have fully acquired the baseline language as children in Sweden, but Martin's earlier and more abrupt onset of English has caused attrition to a higher extent. Among the Heritage speakers that were born in America, early onset of English appears to matter to some extent: Amos and Norman show more deviations than e.g., Arthur, Albert, and Konrad. Amos has also gone the longest without speaking or hearing Swedish. For these speakers, early English generally correlates with less Swedish later in life. It is worth noting that Norman has a high frequency of deviations, despite the fact that he was a monolingual until the age of five, and despite the fact that he is married to Annie (who is clearly more used to speaking Swedish). We return to this briefly in Section Attrition, Acquisition, and Relearning below.

Now, the difference between the speakers is not only how frequent the deviations are, or to what extent the complexity of the noun phrase matters. The speakers also show different patterns with respect to which forms show deviations. We will therefore look a bit closer at the types of deviations.

One weakness with the data is that the different genders are not equally represented: Neuter is much less common than common gender. For instance, Konrad has no complex noun phrases in the neuter, and he produces 74 simple noun phrases with common gender, but only seven with expected neuter. His only deviation involves common gender for neuter. Moreover, it might appear from **Table 7** that Norman has little difficulty with simple noun phrases, but all of the examples involve determiners and quantifiers with common gender morphology; the non-target example is the only case with a neuter noun. Also in the complex noun phrases, the deviating examples have common gender forms for neuter. Here, there are, however, also target-like examples with neuter. The fact that Norman produces some examples of neuter morphology suggests that he has some knowledge of the distinction, but our data is unfortunately not conclusive as to whether his use of gender morphology is systematic or not.

From the limited data, we can note that Norman seems to have a tendency to generalize the default[23]. This is also true for a couple of other speakers. All non-target-like examples produced by Martin involve common gender instead of neuter (cf. 16 above). Mostly, it is the indefinite determiner that deviates [*en* for *ett* as in both (16a) and (16b)], but he also has a couple of examples of deviating adjectives [as in (16b)]. Theodor's deviations, too, all involve common gender for expected neuter. Both adjectives and determiners deviate, and to an equal extent. Deviating examples from Theodor are given in (18).

(Heritage Swedish)

(18)   a. *en   flyg*   (target *ett* an.N.SG.INDEF)
          an.C.SG.INDEF. airplane.N.SG
          'an airplane' (Theodore)

       b. *en   hög   skratt*
          a.C.SG.INDEF. high.C.SG.INDEF. laughter.N.SG
          'a high laughter' (Theodore)
          (target *ett* an.N.SG.INDEF *högt* high.N.SG.INDEF)

---

[23]The conclusion is strengthened by examples with predicatives, where common gender is also overgeneralized, and where we, in fact, find no cases with neuter morphology.

In fact, Amos seems to be the only speaker that sometimes uses neuter for common gender. All of these examples involve the neuter definite determiner, *det*, as in (19)[24].

(19) a. *det*   *yngste*
         the.N.SG.DEF youngest
           son (target: *den* 'the'.C.SG.DEF) (Heritage Swedish)
           son.C.SG
         'the youngest son' (Amos)

While Amos's adjectival inflection also deviates from the baseline in a few cases, 7 of his 12 deviations are of the type in (19), involving the prenominal definite determiner. In fact, Amos has no examples of a common gender definite determiner. It seems then that the determiner *det* is unmarked for gender, for this speaker. We can also note that Amos uses the definite determiner *det* in simplex noun phrases (four examples), combining it with the definiteness suffix (20). Here, the target would involve the simple noun in definite form, without the pre-posed determiner. (Note also that double definiteness is missing in (19) above; Double definiteness would here be required in the baseline.)

(20) *det*   *båt-en*
        the.N.SG.DEF boat. C.SG.DEF
            (target: *den*   'the'.C.SG.DEF) (Heritage Swedish)
        'the boat' (Amos)

The pattern with *det* for expected common gender *den* is not found with the other Swedish speakers (but we saw that this happens quite often with the Norwegian heritage language speakers in complex noun phrases). For other speakers, a common gender determiner is used for a neuter determiner, as in the examples in (18a) and (21), but notice that these are indefinite, and hence not competitors with the neuter definite determiner. There are also target-like uses of the common gender determiner, as in the examples in (22).

---

[24]There are also examples of neuter for common gender in Amos' written production, and these examples are of a different type. They have neuter also on adjectives and other types of determiners than the preposed definite article:

(i) a. *en*   *rolig-t*   *jultid*
         a.C.SG.INDEF fun.N.SG.INDEF christmas.time.
             (target: *rolig* 'fun'. C.SG.INDEF)
             C.SG.INDEF
         'a fun Christmas time' (Amos)

     b. *ett*   *ny-tt*   *flickvän*
         a. N.SG.INDEF new.N.SG.INDEF girlfriend. C.SG.INDEF
         'a new girlfriend' (Amos)
             (target: en 'a'.C.SG.INDEF, ny 'new'. C.SG.INDEF)

Amos does not otherwise treat vän 'friend' as a neuter noun (in his written or oral production); cf. (ii) which shows target-like agreement.

(ii) *en*   *vän*
      a.C.SG.INDEF friend. C.SG.INDEF
          'a friend' (Amos)

(21)   (Heritage Swedish)
  a. *en*   *sjukhus* (target: *et* 'a'.N.SG.DEF)
      a.C.SG.INDEF hospital.N.SG.INDEF
      'a hospital' (Arthur)

  b. *en*   *barn* (target: *et* 'a'. n.sg.def)
      a.C.SG.INDEF child.N.SG.INDEF
      'a child' (Arthur)

(22)   (Heritage Swedish)

    den   *lilla*   *byn*
    the.C.SG.INDEF small.DEF village.C.SG.DEF
    'the small village' (Martin)

Summary on Heritage Scandinavian

The results show, first, that gender is in place in the overall majority of speakers. This is obvious by looking at the Heritage Norwegian CANS corpus, which has only 12% deviant agreement. Second, there are differences in the frequency of deviations in agreement: some speakers show a high frequency of deviations, other speakers have few or no deviant agreement forms. At least to some extent, the frequency of deviations correlate with the speakers' use of the heritage language after childhood, rather than with the context of acquisition. This is particularly clear since one of the immigrant speakers behave more like the American-born heritage speakers than like the other immigrant speaker. At least for this speaker, attrition rather than incomplete acquisition has affected his production of agreement morphology. Early bilinguals also show more deviations than those that were monolingual Scandinavian speakers until school start. Third, several of the speakers appear to have more problems with complex noun phases than simple ones. It seems that the linear complexity of the noun phrase itself can be a factor behind the deviations. We return to complexity in Section Gender, Agreement, and Declension Class. Fourth, as far as we can observe, the deviations belong to the agreement domain, and gender assignment has not developed into only a declension class. Support for this can be found in the fact that there are no deviations in post-nominal possessives (or in the use of the definiteness suffix), but some deviations in the form of pre-nominal determiners. The fact that agreement is also often in place argues against an analysis in terms of loss of gender. Fifth, the data reviewed so far show no clear tendency of overgeneralization of the masculine in Norwegian, but some Swedish speakers seem to have a tendency of overuse of common gender forms (default). Sixth, there seems to be one form that is overused in both languages (creating deviant agreement), viz. *det* 'the'.N.SG.DEF (see further Section Determiners and Adjectives), but in Norwegian this mostly happens in complex noun phrases. Neuter is not otherwise overused in this way. Seventh, nothing in the data suggests that a three-gender system is by itself more vulnerable than a two-gender system, or that feminine gender is particularly vulnerable. Among the Norwegian speakers with the most non-target forms, one has a two-gender system, one a three-gender system. Finally, and importantly, for one phenomenon we do not find inter-individual variation: With a single exception, the form of the definiteness suffix is target-like.

This will be discussed further in Section Attrition, Acquisition, and Relearning.

RESULTS AND DISCUSSION

The results of our study of Heritage Norwegian and Swedish show that with respect to declension (considering the definiteness suffix) there is no variation. In the production of agreement morphology, on the other hand, the speakers show partly different patterns. In this section, we discuss the patterns further. In section Determiners and Adjectives, we look at the difference between determiners and adjectives. Section Morphological Form and Type of Gender System is concerned with the morphology that shows up in the deviant cases in the three-gender system and the two-gender system. In Section Gender, Agreement, and Declension Class, we discuss the definiteness suffix and the question whether gender is on its way of being reduced to declension class. Section Attrition, Acquisition, and Relearning briefly comments on the issue of attrition vs. acquisition in Heritage Scandinavian.

Determiners and Adjectives

As noted in Section Determiners, Adjectives, and Agreement above, in many varieties that have a three-gender system, the feminine is visible on determiners, but rarely on adjectives (the Norwegian adjective *lita* 'little'.f.sg.indef is one of very few adjectives that have a distinct feminine inflectional form). Even so, the Heritage Norwegian data reveals no difficulties regarding this gender, and the feminine determiners are generally used where they should be used according to the baseline.

However, we noted above that some speakers have a different agreement pattern in determiners than in adjectives, even disregarding the feminine. This is clear with the Heritage Swedish speaker Amos, who consistently uses the definite determiner *det* in both neuter and common gender contexts. We noted similar cases in Norwegian. An example of each is repeated from (20) and (10), respectively.

(23) a. *det*   *båten*   (Heritage Swedish)
           the/that.N.SG.DEF boat.c.sg.def
           'the boat' (Amos)
           (target: *den* 'the'.C.SG.DEF)

      b. *det*   *siste*   *plass*   (Heritage Norwegian)
           the.N.SG.DEF last.DEF place.M.SG.INDEF
           'the last place' (chicago_IL_01gk)
            (target: *den*. M.SG.DEF)

From the examples in Section A Closer Look at Two of the Heritage Norwegian Speakers on the two Norwegians and those of American-Swedish Amos, it is clear that substituting other determiners with the neuter *det* 'the'.n.sg.def is a tendency for some of the speakers (but not all). The fact that this determiner in both languages has the pronunciation /de/ means that it is close both in form and meaning to its English counterpart /ðə/. For Amos, it also clearly lacks gender features, like its English counterpart. This possibly is also compatible with the

data from Daisy and Elsa. The indefinite neuter determiner *et/ett* 'a' is not overused in this way, cf. (24). (24c, d) are repeated from (10).

(24) a. *en*   *intressant tur*
           an.C.SG.INDEF interesting trip.C.SG.INDEF
               (Heritage Swedish)
           'an interesting trip' (Amos)

      b. *ett*   *par*   *gånger*   (Heritage Swedish)
           a.N.SG. couple.N.SG. time.PL
           'a couple of times' (Norman)

      c. *en*   *butikk*
           a.M.SG.INDEF shop.M.SG.INDEF
               (Heritage Norwegian)
           'a shop' (chicago_IL_01gk)

      d. *et*   *par*   *koner*
           a.N.SG.INDEF couple.N.SG.INDEF wives
               (Heritage Norwegian)
           'a couple of wives' (chicago_IL_01gk)

This suggests that the overuse of *det* is an effect of phonological similarity, and thus a transfer of the features of a similar functional lexical item in English. We know from other studies of Heritage Scandinavian that functional vocabulary is often affected by transfer (cf. e.g., Larsson et al., 2015). Lexical convergence due to phonological and syntactic similarity is in fact well-known in multilingual settings (Matras, 2009), and has been applied by Annear and Speth (2015) to understand some features of the Heritage Norwegian lexicon. We saw above that the determiner *det* for these speakers sometimes also occurs in simple noun phrases, where it would not be possible in the baseline. A reviewer points out that this could be an additional argument for transfer from English, since English *the* (unlike the No./Sw. article *det*) is used also with simple nouns. The non-neuter definite article *den* does not occur in simple noun phrases in our data (nor in the baseline). Among the Swedish speakers, Amos is the only one who uses *det* with common gender nouns, and he is also the only one who has unstressed *det* in simple noun phrases[25]. We can observe that transfer is more evident in Amos than in the two Norwegian speakers Daisy and Elsa. Unlike Amos, the latter two have some occurrences of the non-neuter article, and the deviations typically occur in complex noun phrases (see Section A Closer Look at Two of the Heritage Norwegian Speakers)[26].

We have divided the results in determiners and adjectives in **Table 8** below; for Norwegian we only include data from Daisy

---

[25]Demonstrative *den/det* (distinguished as it carries stress) is however used in simple noun phrases, as in the baseline language.

[26]A reviewer asks whether the determiner *det* (which in the baseline only occurs in complex noun phrases) does not interfer with the complexity factor. Since only one of the Swedish speakers shows this pattern, and since he uses *det* in both simple and complex noun phrases (to what appears to be an equal extent), this does not affect the overall difference between complex and simple noun phrases. The same seems to be true for Norwegian, where the extended *det* is only present with a limited number of speakers. As noted, above, also complexity appears to matter to a varying extent for different speakers.

TABLE 8 | Non-target forms of all pre-nominal determiners and adjectives in Heritage Norwegian and Swedish.

	Non-target agreement in pre-nominal determiners	Non-target agreement in adjectives
Heritage Norwegian (Daisy and Elsa)	13/89 (15%)	1/15 (7%)
Heritage Swedish	29/318 (9%)	8/72 (11%)
Total	42/407 **(10%)**	9/87 **(10%)**

TABLE 9 | Gender forms in non-target agreement contexts in Heritage Norwegian (two speakers) and Swedish.

	Non-target neuter	Non-target common gender/masculine	Non-target feminine
Heritage Norwegian	7/14 (50%)	4/14 (29%)	3/14 (21%)[28]
Heritage Swedish	7/37 (19%)	30/37 (81%)	(Not applicable)
Total	14/51 **(27%)**	34/51 **(67%)**	3/51 **(6%)**

Feminine inflection generally only on determiners.

and Elsa. Overall the frequency of deviations are very similar in the two cases, even if the examples with overgeneralized *det* are included here.

From **Table 8**, it appears that the Norwegian and Swedish heritage language speakers are maximally different from each other, for while it is the determiners that present the highest amount of non-target forms amongst the former, it is the adjectives that pose the biggest problems for the latter[27]. This difference can be explained by one of the few syntactic differences between the two languages. While Norwegian possessives are generally post-nominal [see examples in (10) above], Swedish ones are, with few exceptions, pre-nominal. If we had included post-nominal possessives, which are always target-like (see **Table 4**), amongst the determiners, the difference would have been much smaller.

It should be noted that in our data set, determiners give more opportunity for deviations than adjectives. This need not mean that adjectives are intrinsically easier, but rather that there are only two gender-relevant adjectival inflections, −Ø and −t, and this difference is only visible in the indefinite paradigm (as in *stor* 'big.M/F.SG.INDEF' and *stort* 'big.N.SG.INDEF'). Thus, in order to get a neuter form of the adjective, three criteria must be satisfied: (1) The noun must be neuter, (2) the noun phrase must be indefinite, and (3) the noun phrase must be singular. Since we also know that neuter nouns are outnumbered by M/F nouns with a factor of 1:4, there will be very few relevant hits in any corpus. In all the indefinite cases of masculine and feminine, the bare stem form would be used. In the definite form and in the plural, there is only one regular form: the −e and −a (Norwegian and Swedish, respectively). The small number of non-target adjectives could be due to this. We therefore do not want to conclude that either determiners or adjectives overall pose more of a challenge to heritage language speakers.

However, from acquisitional studies, we might have expected the heritage speakers to show more difficulty with determiners. In studies of the acquisition of agreement, it has sometimes been noted that determiners cause more difficulty than adjectives. In a study of young (2;7-3;3) Norwegian monolinguals and Norwegian-English bilinguals, Rodina and Westergaard (2013) show that determiners can be unspecified for gender. Other studies, too, have shown that the inflection of determiners might be particularly difficult (see e.g., Cornips and Hulk, 2008 on

Dutch). From previous work on Heritage Scandinavian (see e.g., Larsson and Johannessen, 2015a,b), we know that our heritage speakers sometimes show the same type of patterns as language learners, and we can attribute this to limited input and incomplete acquisition (cf. the introduction). While this might be the source of transfer of functional vocabulary (as for *det* above), it cannot account for the other deviations. As noted, the Swedish immigrant speaker Martin shows deviations, although he presumably acquired Swedish fully in Sweden, before the time of emigration. If, as we argue, the deviations are rather due to attrition and processing difficulty in the adult speakers, it is less clear that the difference between agreeing forms of functional and lexical vocabulary would matter.

Morphological Form and Type of Gender System

For determiners, we saw that Amos and some speakers in the Norwegian corpus used the neuter form *det* /de/ also in common gender/masculine contexts. In other cases, we could observe that the deviant forms most often involved generalization of what we have taken to be the default (masculine or common gender). This pattern appeared to be stronger for Heritage Swedish than for Norwegian. In **Table 9** below, we give the number of non-target uses of neuter, masculine/common gender, and for Norwegian the feminine (which is only marked on determiners).

Non-target neuter can be found in 50% of the Norwegian non-target forms, and 19% of the Swedish ones. These are substantial numbers, but at the same time the difference between the two languages is quite big, and it also turns out that the difference between the speakers is substantial. In fact, all of the Swedish examples where neuter is used for common gender come from the same speaker, Amos. All but one of the Norwegian examples come from Daisy. This suggests that there are individual strategies that are not shared by all the speakers. The loan-transfer of the English determiner *the* to the Norwegian determiner *det* based on similarities in phonological form and syntactic function is one such strategy, as noted above.

Given the arguments for the masculine as a default gender, based on, *inter alia*, the frequency of masculine nouns and their basic phonological form (see Section Gender), it is to be expected that the masculine gender is overused relative to the baseline by heritage language speakers. In the study by Rodina and Westergaard (2013) on monolingual and

---

[27]It should be mentioned that the two speakers do produce several non-target adjectival inflections, but these are related to definiteness and number, and not gender.

[28]But recall that from Section A Closer Look at Two of the Heritage Norwegian speakers that two of the determiners categorized in this cell might actually be target neuter, so this will not be discussed further.

bilingual acquisition, for instance, the majority of errors are overgeneralization of the masculine. Rodina and Westergaard suggest that frequency of forms might be a relevant factor in the acquisition of gender—children overgeneralize the most frequent form (the masculine) in the input. We have not found a clear tendency for all speakers to generalize the masculine, but there is a substantial number of examples (81% in Heritage Swedish and 29% in Heritage Norwegian). Since, the generalization affects determiners as well as adjectives (which take -Ø in the M/F), it cannot exclusively be accounted for in terms of loss of morphological marking.

We have seen several strategies amongst the speakers, but we have not seen anything that suggests a particular pattern based on whether a speaker has a two- or three-gender system. The non-target forms are not completely random, but are in general related to (1) lexical convergence for the neuter singular definite determiner, and (2) overgeneralization of the default masculine/common gender.

Gender, Agreement, and Declension Class

There is, as noted, variation in the production of agreement morphology in Heritage Scandinavian, but there is no corresponding variation in the form of the definiteness suffix. The use of the definiteness suffix is target-like (with one single exception). As noted, similar patterns have been observed for child language. For instance, Rodina and Westergaard (2013) show that in child language, agreement is more error-prone than noun inflection. It has sometimes been suggested that the definiteness suffix is rote learned, i.e., learned as a chunk together with the noun (see e.g., Andersson, 1992, p. 183 and cf. Bohnacker, 2003 for a different view). While this cannot be completely ruled out for individual items, or perhaps for individual speakers (e.g., Amos, who combines the definiteness suffix with a pre-posed definite determiner), we do not see that rote learning can fully account for the target-like definiteness suffix. Firstly, speakers clearly alternate between indefinite and definite forms: Even Amos produces both forms like *hus* 'house.INDEF' and *hus-et* 'house.N.SG.DEF'. Definite forms sometimes occur with what appears to be productive formations, and in code-switching contexts. For instance, Norman uses the form *korporejt-en* 'corporation.C.SG.DEF', and Daisy *river-en* 'river.M.SG.DEF'. Secondly, no speaker attaches plural morphology to the definite form, but the speakers always correctly place plural morphology closer to the root than definiteness morphology (e.g., *kusin-er-na* 'the.cousins.PL.DEF', not *kusin-en-er*, cf. Bohnacker, 2003). One would have expected at least some examples of plural morphology following a singular definite noun if the latter were treated as a chunk. There is therefore at least some tentative evidence that the speakers do analyse the definite forms of nouns, and do not simply treat them as chunks. Thirdly, and importantly in the present context, the definiteness suffix seems to be acquired early by monolingual children in Norway and Sweden (cf. Bohnacker, 2003; Rodina and Westergaard, 2013 and references there), and it is likely to have been fully acquired by the heritage speakers in the present study; this is also what our results suggest. With respect to gender, the Scandinavian heritage speakers thus have a clear

advantage over L2 learners, for whom gender assignment is difficult (see e.g., Andersson, 1992; cf. Montrul et al., 2012, 2014). Moreover, it seems that the definiteness suffix is not affected by attrition. If the heritage speakers can access the noun, they can also access its declension class and (perhaps) its gender.

As far as we can see, there are at least two ways of interpreting the difference in behavior with respect to the suffix and agreement morphology, taking gender to be a lexical category of the noun that is visible in agreement morphology on adjectives and determiners. One possibility is that the gender system is unstable in Heritage Scandinavian, and that what we have interpreted as gender in the definiteness suffix has been (or is on its way to be) reduced to pure declension class. We know from other studies of heritage language that gender systems can be vulnerable (see e.g., Montrul et al., 2008; Polinsky, 2008). The other possibility is that gender is in fact stable in Heritage Scandinavian, and that the variation is more superficial, with the cases of non-target-like agreement being production errors. In the former case, we expect the deviations in agreement to be systematic, and possibly follow the patterns we know from historical changes in the gender system of Scandinavian (e.g., that feminine disappears, or that gender is maintained longer in determiners than adjectives). In the latter case, we expect the type of task and the processing difficulty to be factors, and we expect the behavior of the heritage speakers to be more inconsistent.

We believe that the gender system in Heritage Scandinavian is overall stable. Firstly, most speakers show no or few deviations in agreement (regardless of whether they have a two- or a three-gender system). Secondly, we cannot see that determiners maintain gender distinctions to a higher degree than adjectives (or the other way around).

Thirdly, we see from **Table 4** in Section Agreement in the Corpus of American Norwegian Speech that complexity is important for the Norwegian heritage speakers. In complex noun phrases Daisy and Elsa have 50% non-target forms, while the number is reduced to 8% for the simpler adjective–noun combinations, and down to 0% in the simplest noun phrases (with post-posed possessives). The Swedish heritage speakers have 16 and 7% for the most complex and the least complex ones (see **Table 7**, Section Gender Agreement in Complex and Simple Noun Phrases). The individual differences are big, and might account for the difference between the two groups. However, for the speakers that show deviations to a higher degree, it seems that the task of applying the same gender morphology to several items in a noun phrase is the biggest problem.

Importantly, complexity here does not necessarily mean structural (syntactic) complexity. Following Julien (2005) and others, we can assume that structures with post-nominal possessives are structurally more complex than structures with pre-nominal possessives (the former involving movement of the head noun). As shown by Anderssen and Westergaard (2010), monolingual Norwegian children seem to prefer the syntactically less complex order with pre-nominal determiner and use it also in contexts where it is not used in the input (where the post-nominal determiner is more frequent). Based on this and other evidence, they argue that structural complexity, rather than frequency determines the path of acquisition. In a later study

(Westergaard and Anderssen, 2015), they show that the same does not necessarily hold among Norwegian heritage speakers, who rather overuse the post-nominal possessive. They conclude that structural complexity is not a factor for attrition in the same way as it is in acquisition.

With respect to agreement morphology, too, prenominal determiners appear to be more difficult than post-nominal possessives. Instead of structural complexity, it seems that the linear distance between agreeing form and head noun matters for our speakers. This suggests that the speakers generally have an intact gender system, which they fail to adhere to in situations that are demanding for their working memory. Myles (1995) makes a similar case for second language acquisition in French, where she suggests that agreement morphology is sensitive to level of embedding. She argues that the degree of automatization in processing is crucial; only when low-level processes (processing in local domains) are automatic, is the short-term memory freed and can deal with higher-level processes (cf. e.g., Pienemann, 1998 and many others on processability).

More recent studies have confirmed Myles (1995) results, but there is some disagreement as to the reason behind the difficulties (see Keating, 2009; Foote, 2011 and references there). For Myles, the relevant factor in second language acquisition is structural complexity. As pointed out above, in our study it rather seems to be linear distance that matters. This is more in line with results like those in Keating (2009), who shows with an eye-tracking experiment that while advanced learners of Spanish have acquired gender distinctions, they are non-native-like by being affected by the distance between nouns and modifiers[29]. The fact that our heritage speakers have not used the language regularly for many years clearly has effects on the burden on their short-term memory and processing abilities. For instance, lexical retrieval is known to become less automatic (and more costly) in attrition (Polinsky, 2008), and gender is clearly tied to lexical retrieval.

Our conclusion is that the underlying Scandinavian gender system is not particularly vulnerable, so that the definiteness suffixes of the nouns, which are always correct, actually signify gender and not just declension class. At the same time we see some signs of vulnerability amongst the most attrited speakers, where we find both lexical convergence of the definite determiner *det* /de/ with English *the* /ðə/, and a (somewhat weak) tendency to generalize the masculine/common gender.

Thus, our result is in contrast with previous studies of other heritage languages, where gender as noted has been shown to be vulnerable (e.g., Polinsky, 2008; Montrul et al., 2008, 2014). For instance, in a study of Spanish gender, Alarcón (2011) concludes that gender assignment (rather than agreement) is particularly sensitive to incomplete acquisition, and Polinsky (2008) observes systematic changes in the gender system of Heritage Russian. For Heritage Scandinavian, on the other hand, gender agreement appears to be more sensitive than gender assignment, and there

are no systematic changes in the gender system (again, if we disregard *det* in individual speakers). The question, then, is why Scandinavian would be different.

Now, we expect that some of the more systematic changes in gender systems in heritage languages are due (to a higher extent) to incomplete acquisition, rather than attrition. In fact, it is possible that the differences between e.g., Spanish and Russian heritage language, on the one hand, and Heritage Scandinavian, on the other, stem from differences in the acquisitional process. As pointed out by Bohnacker (2003), studies of L1 acquisition of Swedish (e.g., Plunkett and Strömqvist, 1992; Andersson, 1992) suggest that gender is acquired early and with ease, in contrast with acquisition of gender in many other languages. This appears to be the case even if Scandinavian gender is largely unpredictable from phonology and semantics. One possible explanation, suggested by Andersson (1994), is that the evidence for the gender of a noun in Swedish is not only found in agreement patterns, but also in the definiteness suffix, which unambiguously signals the gender of the noun (cf. the discussion above). Thus, the acquisition of gender can go hand in hand with the acquisition of declension class. We thus hypothesize that the vulnerability of gender in other heritage languages could stem from incomplete acquisition, though the deviations we have noted in the present study are largely a consequence of attrition.

However, if this is on the right track, and the cross-linguistic differences correlate with differences in the acquisition of gender (due to the varying evidence for gender), one might expect a clear difference between Heritage Swedish and Heritage Norwegian, contrary to what we find. Recent studies of acquisition of gender in Norwegian (Gagliardi, 2012; Rodina and Westergaard, 2013, 2015) have argued that gender assignment (not agreement) is in fact difficult for children. Particularly the feminine gender appears vulnerable, and deviations from adult language might persist well into school age. (There are also deviations in the use of neuter, but to a lesser degree; Rodina and Westergaard, 2015, p. 176). It is possible that this difference between Swedish and Norwegian is due to the different gender systems (two vs. three genders), but it is also likely that the linguistic situation in Norway has something to say. Norwegian children are typically exposed to more than one gender system, since one of the standard varieties (Bokmål) can have a two-gender system (or perhaps a few remnants of the feminine gender), whereas most dialects have three genders. On the basis of data from different generations of speakers of the Tromsø dialect, Rodina and Westergaard (2015) in fact argue that the gender system is changing, and that the feminine gender appears to be on its way out. In addition to the bi-dialectal situation, they point to independent changes that lead to the loss of some of the morpho-phonological cues for the feminine (2015:181). Moreover, in the adjectival inflection, the feminine gender is as noted only rarely distinguished from the masculine. Crucially, the situation for the Norwegian heritage speakers is quite different, since there is no influence from the standard language (cf. Johannessen and Laake, forthcoming). Instead, the heritage speakers generally only speak and understand their own dialect, and they have no knowledge of the written language.

---

[29] Also other studies show that there is some evidence for so-called shallow processing in second language learners (see e.g., Sorace, 2006 for discussion). In the context of gender agreement Keating (2009) operationalizes shallow processing "in terms of the distance that separates agreeing constituents."

It is therefore possible that the situation for the Norwegian heritage speakers is more similar to that of the Swedish speakers. However, some caution is required in the interpretation of the results, since the studies of acquisition in Norway focus on particular dialects, since the individual variation among the heritage speakers is considerable, and since the different studies employ partly different methodologies. Additional work is clearly required.

Attrition, Acquisition, and Relearning

The deviations in agreement seem to be a consequence of language attrition at the level of the individual. The speakers in the present study are old (with the exception of Martin, 55 years old at the time of the recording) and have not spoken Scandinavian daily for many years. The number of deviations does not necessarily correlate with acquisitional context. Most of the speakers that show no deviations are born in the USA, and one of our two immigrant speakers show agreement deviations.

At the same time, it seems that the onset of the dominant language (English) is an important factor. As noted for Swedish, the speakers with the highest number of deviations were all early bilinguals. This is also true for the Norwegian speaker Daisy. A more typical pattern is acquisition of the majority language at school start.

From the perspective of early bilingualism and language use later in life, it is perhaps surprising that the heritage speakers Norman (Swedish) and Elsa (Norwegian) show a relatively high number of deviations. Norman reports that he was monolingual until the age of five, and he is married to a Swedish immigrant. He reported that he started speaking English with his parents before school start, in order to prepare for school, and it seems that his connections to the Swedish community had mostly been through his wife, who he met when he was around 20 years old. Elsa also has had a lot of contact with Norway at an adult age, given that two of her children have married there. It is possible, we think, that the deviations for both Elsa and Norman can to some extent be a consequence of language loss followed by relearning. It is possible that relearning makes the heritage language less native-like, and perhaps even L3-like. Polinsky (2015a) shows that, in an environment of instruction of the written baseline language, heritage speakers do outperform L2 learners in the perception and production of phonology. However, when learning grammatical features, heritage speakers were outperformed. Viswanath (2013, p. 39) further shows that heritage speakers over-regularize forms in a learning context. It seems, therefore, that relearning does not necessarily improve the heritage speakers' competence. The fact that Norman and Elsa have had close encounters with the European Scandinavian varieties as adults might have had the same effect as relearning in a formal context.

One difficulty in their relearning is that the relearnt language (the modern Scandinavian dialects/standards) is in fact substantially different from their original heritage language (compare Polinsky, 2015a). Further studies on heritage language and relearning are required.

CONCLUSION

In this paper, we have looked at agreement in the noun phrase. The Scandinavian noun phrase clearly poses several difficulties in language acquisition, including double definiteness marking (Anderssen, 2012, p. 4), agreement and gender assignment to the noun, and it has previously been shown to be affected in attrition.

To sum up our investigation, the results show the following. First, gender is in place in the overall majority of all our speakers. Among the 34 speakers in the Heritage Norwegian CANS corpus, there is only 12% deviant agreement with respect to the baseline, and among the eight Heritage Swedish speakers only 10% deviance. Second, there is considerable variation among speakers. Nearly 60% of the Norwegian speakers show no deviant forms, while a few speakers show a considerable number. Of the Swedish Heritage and Immigrant speakers, five speakers have deviance in 0–9% of their total number of noun phrases, while three have 19–28%. Third, we see that the complexity of the noun phrase matters: The two most attrited Norwegian speakers have 50% deviance in complex noun phrases, and 0–8% in various kinds of simple noun phrases. The Swedish speakers have 16% deviance in complex noun phrases and only 7% in simple noun phrases. This suggests that the deviations are due to processing difficulty. Fourth, the deviations belong to the agreement domain, and not in the definiteness suffix. It can be concluded that gender assignment is largely in place, and that the definiteness suffix has not developed into a marker of just declension class. Support for this can be found in the fact that there are no deviations in the Norwegian post-nominal possessives (or in the use of the definiteness suffix), and only a few deviations in the form of pre-nominal determiners, in both languages. Fifth, the data show that there is a tendency to overgeneralize the masculine (which is the default gender), but there is also one particular neuter form which is overused (creating deviant agreement), probably due to its similarity in form and meaning to its English counterpart, viz. *det* /de/, 'the'.N.SG.DEF, from English, similar to *the* /ðə/ DEF. No other forms of neuter are overused in this way, clearly showing that this is an effect of lexical convergence. Finally, nothing in the data suggests that a three-gender system is by itself more vulnerable than a two-gender system, or that feminine gender is particularly vulnerable. Among the Norwegian speakers with the most non-target forms, one has a two-gender system, one a three-gender system. The patterns we can observe in Norwegian three-gender speakers are also found among the Swedish two-gender speakers.

We would like to point out that the data we have used come from a variety of speakers (Norwegian and Swedish, heritage and immigrant) and sources (in depth studies of interviews and automatic counts of a large corpus), and that this has given us the possibility to investigate both general patterns and inter-speaker variation, and to explore different types of explanations. Going back to the factors that have previously been shown to affect the properties of Heritage Scandinavian, we can note the particular acquisitional context of the American-born heritage speakers do not necessarily affect gender agreement. Moreover, there is overall very little evidence of transfer from English (with a single exception). We do not see a general simplification of

the gender system across Heritage Scandinavian. Instead, we have considerable inter-individual variation. We have observed that one immigrant speaker show more deviations than many of the American-born heritage speakers. For that reason, among others, we have wanted to argue that the deviations we find are due to attrition. Given previous studies, this is perhaps not surprising: Morphology has been shown to be sensitive in attrition. As expected, it appears that the time of onset of English matters for the degree of attrition, in combination with the use of the L1 later in life. However, several factors are clearly intertwined, and they call for further study; for some speakers, relearning might be involved, as well. For some individuals, there are specific deviations that might also be due to reanalysis in the first language acquisition. This would be one way of accounting for the lexical convergence of one determiner.

Notably, this change is restricted to a single functional word.

ACKNOWLEDGMENTS

REFERENCES

Alarcón, I. (2011). Spanish gender agreement under complete and incomplete acquisition: early and late bilinguals' linguistic behavior within the noun phrase. *Bilingual. Lang. Cogn.* 14, 332–350. doi: 10.1017/S1366728910000222

Anderssen, M., and Westergaard, M. (2010). Frequency and economy in the acquisition of variable word order. *Lingua* 120, 2569–2588. doi: 10.1016/j.lingua.2010.06.006

Anderssen, M. (2012). A spanning approach to the acquisition of double definiteness in Norwegian. Iberia: an International. *J. Theor. Linguist.* 4.1, 1–34.

Andersson, A. (1992). *Second Language Learner's Acquisition of Grammatical Gender in Swedish.* Diss. Department of Linguistics: Gothenburg University.

Andersson, A. (1994). "On the Acquisition of Gender in Swedish," in *Proceedings of the XIVth Scandinavian Conference of Linguistics and the VIIIth Conference of Nordic and General Linguistics: Special Session on Language Development*, ed Sven Strömqvist (Department of Linguistics; Gothenburg University).

Andréasson, M., Larsson, I., Lyngfelt, B., Nilsson, J., and Tingsell, S. (2013). "I jakt på amerikasvenskan," in *Svenskans Beskrivning 32*, ed B. Bihl, J. Eriksson, P. Andersson, and L. Lötmarker (Karlstad: Karlstads universitet), 82–91.

Annear, L., and Speth, K. (2015). "Maintaining a multilingual repertoire: lexical change in American Norwegian," in *Germanic Heritage Languages in North America Acquisition, Attrition and Change*, eds J. B. Johannessen and J. C. Salmons (Amsterdam: John Benjamins), 201–216.

Baker, M. (2008). *The Syntax of Agreement and Concord.* Cambridge: Cambridge University Press.

Benmamoun, E., Montrul, S., and Polinsky, M. (2013). Heritage languages and their speakers: opportunities and challenges for linguistics. *Theor. Linguist.* 39, 129–181. doi: 10.1515/tl-2013-0009

Bohnacker, U. (2003). "Nominal phrases," in *The Acquisition of Swedish Grammar*, eds G. Josefsson, C. Platzack, and G. Håkansson (Amsterdam: John Benjamins), 195–260.

Corbett, G. G. (1991). *Gender.* Cambridge: Cambridge University Press.

Cornips, L., and Hulk, A. (2008). Factors of success and failure in the acquisition of grammatical gender in Dutch. *Sec. Lang. Res.* 24, 267–369. doi: 10.1177/0267658308090182

Dahl, Ö. (2015). *Grammaticalization in the North: Noun Phrase Morphosyntax in Scandinavian Vernaculars.* Berlin: Language Science Press.

Davidson, H. (1990). *Han hon den.* Lund: Lund University Press.

Delsing, L.-O. (1993). *The Internal Structure of Noun Phrases in the Scandinavian Languages. A Comparative Study.* Department of Scandinavian Languages; University of Lund.

Eide, K. M., and Hjelde, A. (2012). "V2 and morphological paradigms in Norwegian varieties spoken in the American Midwest," in *Third Workshop of Immigrant Languages in America*, State College, PA: Penn State University.

Enger, H.-O. (2004). On the relationship between gender and declension: a diachronic perspective from Norwegian. *Stud. Lang.* 28, 51–82. doi: 10.1075/sl.28.1.03eng

Enger, H.-O. (2014). Reinforcement in inflection classes: two cues may be better than one. *Word Struct.* 7, 153–181. doi: 10.3366/word.2014.0064

Faarlund, J. T., Svein, L., and Ivar Vannebo, K. (1997). *Norsk Referansegrammatikk.* Oslo: Universitetsforlaget.

Foote, R. (2011). Integrated knowledge of agreement in early and late English-Spanish bilinguals. *Appl. Psycholinguist.* 32, 187–220. doi: 10.1017/S0142716410000342

Fretheim, T. (1976/1985). "Er bokmålet tvekjønnet eller trekjønnet?" in *Morfologi/Morphology*, eds i Ernst Håkon Jahr og Ove Lorentz (Oslo: Novus), 99–101.

Gagliardi, A. (2012). *The Fundamentals of Language Acquisition.* Baltimore, MD: University of Maryland dissertation.

Håkansson, G. (1995). Syntax and attrition: a study of five bilingual expatriate Swedes. *Int. J. Appl. Linguist* 5, 153–171.

Hasselmo, N. (1974). *Amerikasvenska. En bok om språkutvecklingen i Svensk-Amerika.* Stockholm; Skrifter utg. av Svenska språknämnden 51.

Haugen, E. (1953). *The Norwegian Language in America. A Study in Bilingual Behaviour.* Philadelphia: University of Pennsylvania Press.

Heggstad, K. (1982). *Norsk frekvensordbok.* Oslo; Universitetsforlaget.

Hjelde, A. (1992). *Trøndsk Talemål i Amerika.* Trondheim; Tapir Forlag.

Johannessen, J. B. (2008). Pronominal psychological demonstratives in Scandinavian: their syntax, semantics and pragmatics. *Nord. J. Linguist.* 31, 161–192. doi: 10.1017/S0332586508001923

Johannessen, J. B., and Laake, S. (2012). Østnorsk som norsk fellesdialekt i Midtvesten. *Nor. Lingvistisk Tidsskr.* 30, 365–380. Available online at: http://urn.nb.no/URN:NBN:no-46109

Johannessen, J. B. and Laake, S. (forthcoming). Norwegian in the American Midwest: a common lexicolect. *J. Lang. Contact.* 10.

Johannessen, J. B., Priestley, J., Hagen, K., Åfarli, T. A., and Vangsnes, Ø. A. (2009). "The Nordic dialect corpus – an advanced research tool," in *Proceedings of the 17th Nordic Conference of Computational Linguistics NODALIDA 2009. NEALT Proceedings Series, Vol. 4*, eds K. Jokinen and E. Bick (Estonia: Tartu University Library; Northern European Association for Language Technology), 73–80. Available online at: http://hdl.handle.net/10062/9206

Johannessen, J. B., and Salmons, J. (2012). Innledning. *Nor. Lingvistisk Tidsskr.* 30, 139–148. Available online at: http://urn.nb.no/URN:NBN:no-46110

Johannessen, J. B., and Salmons, J. C. (eds). (2015). *Germanic Heritage Languages in North America Acquisition, Attrition and Change.* Amsterdam: John Benjamins.

Johannessen, J. B. (2012). "Han professoren og sånn festskrift – om modale determinativer," in *Grammatikk, Bruk og Norm. Festskrift til Svein Lie på 70-Årsdagen*, eds H.-V. Enger, Jan Terje F, and V. Kjell Ivar (Oslo: Novus), 151–172.

Johannessen, J. B. (2015a). "Attrition in an American Norwegian heritage language speaker," in *Germanic Heritage Languages in North America Acquisition, Attrition and Change*, eds J. B. Johannessen and J. C. Salmons (Amsterdam: John Benjamins), 46–71.

Johannessen, J. B. (2015b). "Corpus of American Norwegian Speech (CANS)," in *Proceedings of the 19th Nordic Conference of Computational Linguistics (NODALIDA 2015),* ed B. Megyesi (Institute of the Lithuanian Language, Lithuania. NEALT Proceedings Series. Linköping University Electronic Press). Available online at: http://urn.nb.no/URN:NBN:no-48231

Josefsson, G. (2009). Pancakes and peas – on apparent disagreement and (null) light verbs in Swedish. *Nor. J. Linguist.* 32, 35–72. doi: 10.1017/S0332586509002030

Julien, M. (2005). *Nominal Phrases from a Scandinavian Perspective.* Amsterdam: John Benjamins Publishing Company.

Källström, R. (1996). "On gender assignment in Swedish," in *The Nordic Languages and Modern Linguistics 9,* ed K. G Ottósson, R. V. Fjeld, and A. Torp (Oslo: Novus Forlag), 151–167.

Källström, R. (2008). "Varför är det så gott om neutrum?," in *Svenskans Beskrivning 29, Förhandlingar vid Tjugonionde Sammankomsten för Svenskans Beskrivning,* eds N. Marianne, S. Björklund, C. Laurén, K. Mård-Miettinen, and N. Pilke (Vasa: Svensk-Österbottniska Samfundet), 179–186.

Keating, G. D. (2009). Sensitivity to violations of gender agreement in native and nonnative Spanish: an eye-movement investigation. *Lang. Learn.* 59, 503–535. doi: 10.1111/j.1467-9922.2009.00516.x

Kupisch, T., Lein, T., Barton, D., Schröder, D. J., Stangen, I., and Stoehr, A. (2014). Acquisition outcomes across domain in adult heritage speakers of French. *J. French Lang. Stud.* 24, 347–376. doi: 10.1017/S0959269513000197

Larsen, A. B., and Stoltz, G. (1911). *Bergens Bymål.* Kristiania; Bymålslaget.

Larsson, I., and Johannessen, J. B. (2015a). "Embedded word order in Heritage Scandinavian," in *The 11th International Conference of Nordic and General Linguistics,* eds H. Martin, M. Jan-Ola, C. Östman, M. R. Mertzlufft, and D. Janet (Berlin: Linguae et Litterae, Mouton de Gruyter), 239–266.

Larsson, I., and Johannessen, J. B. (2015b). "Incomplete acquisition and verb placement in Heritage Scandinavian," in: *Moribund Germanic Heritage Languages in North America: Theoretical Perspectives and Empirical Findings,* eds R. S. Putnam and T. Michael (Leiden: Brill Academic Publishers), 153–189.

Larsson, I., Tingsell, S., and Andréasson, M. (2015). "Variation and Change in American Swedish," in *Germanic Heritage Languages in North America Acquisition, Attrition and Change,* eds J. B. Johannessen and J. C. Salmons (Amsterdam: John Benjamins), 359–388.

Leu, T. (2015). *The Architecture of Determiners.* Oxford: Oxford University Press.

Lødrup, H. (2011). Hvor mange genus er det i Oslo-dialekten? *Maal og Minne* 103, 120–136. Available online at: http://ojs.novus.no/index.php/MOM/article/view/522

Lohndal, T., and Westergaard, M. (2014). "Gender attrition in American Norwegian Heritage Language," in *Presentation at 5th Annual Workshop on Immigrant Languages in the Americas* (Los Angeles: University of California).

Matras, Y. (2009). *Language Contact.* New York, NY: Cambridge University Press.

Montrul, S., Davidson, J., de la Fuente, I., and Foote, R. (2014). Early language experience facilitates the processing of gender agreement in Spanish heritage speakers. *Bilingualism* 17, 118–138. doi: 10.1017/S1366728913000114

Montrul, S., de la Fuente, I., Davidson, J., and Foote, R. (2012). The role of experience in the acquisition and production of diminutives and gender in Spanish: Evidence from L2 learners and heritage speakers. *Sec. Lang. Res.* 29, 87–118. doi: 10.1177/0267658312458268

Montrul, S., Foote, R., and Perpinán, S. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: the effect of age and context of acquisition. *Lang. Learn.*58, 503–553. doi: 10.1111/j.1467-9922.2008.00449.x

Montrul, S. (2008). *Incomplete Acquisition in Bilingualism. Re-Examining the Age Factor.* Amsterdam: John Benjamins.

Myles, F. (1995). Interaction between linguistic theory and language processing in SLA. *Sec. Lang. Res.* 11, 235–266.

Nygård, M., and Åfarli, T. A. (2013). *Genus i Amerikanorsk og Strukturen til Gensutilordning. RISS: Magasin for Studentar og tilsette Ved,* Vol. 2. Institutt for nordistikk og litteraturvitskap; Universitetet i Trondheim.

Pascual y Cabo, D., Lingwall, A., and Rothman, J. (2012). "Applying the interface hypothesis to heritage speaker acquisition: evidence from Spanish mood," in *BUCLD 36: Proceedings of the 36th Annual Boston University Conference on Language Development.* (Sommerville, MA: Cascadilla Press), 437–438.

Pascual y Cabo, D., and Rothman, J. (2012). The (Il)logical problem of heritage speaker Bilingualism and incomplete acquisition. *Appl. Linguist.* 33, 450–455. doi: 10.1093/applin/ams037

Pienemann, M., and Håkansson, G. (1999). A unified approach towards the development of Swedish as L2: a processability account. *Stud. Sec. Lang. Acquisition* 21, 383–420.

Pienemann, M. (1998). *Language Processing and Second Language Development: Processability Theory.* Amsterdam: John Benjamins.

Plunkett, K., and Strömqvist, S. (1992). "The acquisition of Scandinavian languages," in *The Crosslinguistic Study of Language Acquisition,* ed D Slobin (Hillsdale, NJ: Lawrence Erlbaum Associates), 457–556.

Polinsky, M. (2008). Gender under incomplete acquisition: heritage speakers' knowledge of noun categorization. *Herit. Lang. J.* 6, 40–71. Available online at: https://dash.harvard.edu/handle/1/3382970

Polinsky, M. (2015a). When L1 becomes an L3: assessing grammatical knowledge in heritage speakers/learners. *Bilingual. Lang. Cogn.* 18, 163–178. doi: 10.1017/S1366728913000667

Polinsky, M. (2015b). Heritage languages and their speakers: state of the field, challenges, perspectives for future work, and methodologies. *Z. Fremdsprachwissenschaft* 26, 7–27. Available online at: http://scholar.harvard.edu/files/mpolinsky/files/heritage_paper_new.pdf?m=1442332868

Rodina, Y., and Westergaard, M. (2013). The Acquisition of gender and declension class in a non-transparent system: monolinguals and bilinguals. *Stud. Linguist.* 67, 47–67. doi: 10.1111/stul.12012

Rodina, Y., and Westergaard, M. (2015). Grammatical Gender in Norwegian: language acquisition and language change. *J. Ger. Linguist.* 27, 145–187. doi: 10.1017/S1470542714000245

Sandøy, H. (1988). "Samsvarsbøying av adjektiv og perfektum partisipp i norske dialektar," in *Nordiske Studiar. Innlegg på den Tredje Nordiske Dialektologkonferansen,* eds Andreas Bjørkum and Arve Borg (Oslo: Universitetsforlaget), 85–117.

Sorace, A. (2004). Native language attrition and developmental instability at the syntax-discourse interface: data, interpretation and methods. *Bilingualism* 7, 143–145. doi: 10.1017/S1366728904001543

Sorace, A. (2006). Possible manifestations of shallow processing in advanced second language speakers. *Appl. Psycholinguist.* 27, 88–91. doi: 10.1017/S0142716406060164

Teleman, U. (1969). On gender in a generative grammar of Swedish. *Stud. Linguist.* 23, 27–67.

Trosterud, T. (2001). Genustilordning i norsk er regelstyrt. *Nor. Lingvistisk Tidsskr.* 19, 29–59.

Trudgill, P. (2013). "Gender maintenance and loss in Totenmålet, English, and other major Germanic varieties," in *In Search of Universal Grammar: From Old Norse to Zoku,* ed L. Terje (Amsterdam: John Benjamins), 77–108.

Vangsnes, Ø. A., Holmberg, A., and Delsing, L- O. (red.). (2003). *Dialektsyntaktiska Studier av den Nordiska Nominalfrasen.* Oslo: Novus forlag.

Viswanath, A. (2013). *Heritage English in Israeli children.* BA Thesis, Harvard University, Cambridge, MA.

Westergaard, M., and Anderssen, M. (2015). "Word order variation in Norwegian possessive constructions: bilingual acquisition and attrition," in *Germanic Heritage Languages in North America Acquisition, Attrition and Change,* eds J. B. Johannessen and J. C. Salmons (Amsterdam: John Benjamins), 21–45.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

CrossMark

Grammatical Gender in American Norwegian Heritage Language: Stability or Attrition?

Terje Lohndal[1,2]* and Marit Westergaard[2,1]

[1] *Department of Language and Literature, Norwegian University of Science and Technology, Trondheim, Norway,*
[2] *Department of Language and Culture, University of Tromsø The Arctic University of Norway, Tromsø, Norway*

This paper investigates possible attrition/change in the gender system of Norwegian heritage language spoken in America. Based on data from 50 speakers in the Corpus of American Norwegian Speech (CANS), we show that the three-gender system is to some extent retained, although considerable overgeneralization of the masculine (the most frequent gender) is attested. This affects both feminine and neuter gender forms, while declension class markers such as the definite suffix remain unaffected. We argue that the gender category is vulnerable due to the lack of transparency of gender assignment in Norwegian. Furthermore, unlike incomplete acquisition, which may result in a somewhat different or reduced gender system, attrition is more likely to lead to general erosion, eventually leading to complete loss of gender.

Keywords: acquisition, American Norwegian, declension class, gender, heritage language

INTRODUCTION

In his seminal study, Corbett (1991, p. 2) states that "[g]ender is the most puzzling of the grammatical categories." It involves the interaction of several components: morphology, syntax, semantics, phonology, as well as knowledge about the real world. Languages also differ in terms of how many (if any) genders they have. This means that gender is a property of language which must be inferred from the input to which both child and adult learners of a language have to be finely attuned.

We follow Hockett (1958, p. 231) in defining gender as follows: "Genders are classes of nouns reflected in the behavior of associated words." This means that gender is expressed as agreement between the noun and other elements in the noun phrase or in the clause and that affixes on the noun expressing e.g., case, number or definiteness are not exponents of gender (Corbett, 1991, p. 146). We refer to the marking on the noun itself as an expression of declension class (cf. Enger, 2004; Enger and Corbett, 2012; see also Kürschner and Nübling, 2011 for a general discussion of the difference between gender and declension class in the Germanic languages). This has an interesting consequence for the definite article in Norwegian, which is a suffix (more on this below). A distinction is also commonly made between gender assignment and gender agreement. Gender assignment is what is typically referred to as an inherent property of the noun, e.g., $bil_{(M)}$ "car" and $hus_{(N)}$ "house," while gender agreement refers to agreement on other targets that is dependent on the gender of the noun, e.g., the indefinite articles and adjectives in $en._M fin._M bil_{(M)}$ "a nice car" and $et._N fint._N hus_{(N)}$ "a nice house".[1] The literature also differentiates between lexical vs. referential

---

[1] We indicate gender on the noun itself in parenthesis and gender agreement on other targets after a period.

gender (Dahl, 2000), or in the terminology of Corbett (1991), syntactic vs. semantic gender. The former refers to the inherent and invariable gender of a noun, e.g., *papa* "daddy" in Russian, which is always masculine, whereas the other refers to cases where gender depends on the referent, e.g., *vrac* "doctor," which may take either feminine or masculine agreement.

In this article, we provide a case study of gender assignment in a population of heritage speakers of Norwegian who have lived their entire lives in America, often without ever visiting Norway. We follow Haugen (1953) in referring to this variety as American Norwegian, and here we study whether the use of gender differs in any way from the traditional use of gender in Norwegian dialects. We are also interested in the nature of possible discrepancies. This will provide important information on how gender systems may change over time, especially in contexts with reduced input and use, and we compare the situation in American Norwegian to heritage Russian spoken in the US. As Polinsky (2008, p. 40) emphasizes, "[s]ince very little is actually known about heritage language speakers, studying different aspects of language structure in this population is important." The current paper contributes to this end in that it provides an additional investigation into the linguistic structure of heritage languages.

The structure of the paper is as follows. In the next section, we introduce gender and its manifestations within the Norwegian noun phrase. We then outline some relevant background from acquisition and heritage contexts, and the following section introduces our research questions, participants, and methodology. We then present our results, followed by a discussion and some concluding remarks.

GENDER AND THE NORWEGIAN NOUN PHRASE

Norwegian dialects traditionally distinguish between three genders: masculine, feminine and neuter. While many languages with gender have reliable morphophonological gender cues, e.g., Spanish or Italian (where a noun ending in *–o* marks masculine and *–a* marks feminine), gender assignment in Norwegian is non-transparent. That is, from just hearing a noun, e.g., *bil* "car," *bok* "book," or *hus* "house," a learner cannot make out its gender. It is only when nouns appear with associated words that the gender can be identified, e.g., the indefinite article, as in *en*.N *bil*(M), *ei.f bok*(F), and *et.n hus*(N). Nevertheless, Trosterud (2001) proposes 43 different assignment rules and argues that they may account for 94% of all nouns in the language. These assignment rules include three general rules, nine morphological rules, three phonological rules, and 28 semantic rules. However, each rule has numerous exceptions, making it less clear if or how this rule-based account could actually predict gender in acquisition situations. Thus, we follow Rodina and Westergaard (2013, 2015a,b) in assuming that the acquisition of gender in Norwegian is opaque and must be learned noun by noun. This makes Norwegian gender a challenging property to acquire in a heritage language situation,

where there is typically reduced input (see O'Grady et al., 2011).

Norwegian has two written standards, *Nynorsk* and *Bokmål*, the latter being by far the dominant one (see Venås, 1993 for more information about the Norwegian language situation). In *Bokmål*, all feminine nouns may take masculine agreement, which means that this written variety may use only two genders, common and neuter. The historical reason for this is that *Bokmål* is a development of the Danish written standard, and in Danish (as well as in Swedish and Dutch) the gender system has been reduced from one that distinguished three genders to one that generally only has two. The three-gender system has generally been retained in spoken Norwegian, in virtually all dialects (except Bergen and parts of Oslo). However, some recent studies indicate that a change from a three-gender system to a two-gender system is underway in the Tromsø dialect (Rodina and Westergaard, 2015a). More about this below.

Norwegian noun phrase syntax is relatively complex, and it has been extensively discussed in the literature; see Delsing (1993), Vangsnes (1999), and Julien (2005). Here we only discuss aspects of the noun phrase that are relevant for gender. Norwegian dialects also differ considerably with respect to the specific morphological marking on nouns. **Table 1** provides an overview of the three-way gender system (based on the written *Bokmål* norm).

Gender in Norwegian is mainly expressed inside the noun phrase (and on predicative adjectives, not discussed in this article). Thus, gender is marked on the indefinite article, e.g., *en* "a.M," *ei* "a.F" and *et* "a.N," and on adjectives, where we find syncretism between M and F forms[2].

TABLE 1 | The traditional three-gender system of Norwegian.

Gender	Masculine	Feminine	Neuter
Indefinite	**en** hest *a horse*	**ei** seng *a bed*	**et** hus *a house*
Definite	hest**en** *horse.DEF*	seng**a** *bed.DEF*	hus**et** *house.DEF*
Double definite	**den** hest**en** *that horse.DEF*	**den** seng**a** *that bed.DEF*	**det** hus**et** *that house.DEF*
Adjective	en **fin** hest *a nice horse*	ei **fin** seng *a nice bed*	et **fint** hus *a nice house*
Possessive	**min** hest/hest**en min** *my horse/horse.DEF my*	**mi** seng/seng**a mi** *my bed/bed.DEF my*	**mitt** hus/hus**et mitt** *my hourse/house.DEF my*

---

[2]There is only one exception to this, the adjective *liten/lita/lite* "small/little," which distinguishes between all three genders. This is illustrated in (i).

(i) a. *en liten gutt*
　　　a.M small.M boy
　　　"a small boy"
　　b. *ei lita jente*
　　　a.F small.F girl
　　　"a small girl"
　　c. *et lite hus*
　　　a.N small.N house
　　　"a small house"

As shown in **Table 1**, the definite article in Norwegian is a suffix, e.g., *hesten* "the horse," *senga* "the bed," *huset* "the house." Some traditional grammars of Norwegian analyze the post-nominal definite suffix as an expression of gender (e.g., Faarlund et al., 1997), mainly because it is derived diachronically from postnominal demonstratives (separate words), which used to be marked for gender. Given our definition in the Introduction, however, these suffixes do not express gender, but should be considered to be declension class markers.

Since the definite suffix is sometimes considered to express gender, also in current work (e.g., Johannessen and Larsson, 2015), it is worth pausing to consider the evidence in favor of suffixes being declension class markers. This view is most prominently articulated by Lødrup (2011), based on a careful investigation of (a variety of) the Oslo dialect, where the feminine gender is argued to have been lost. The main piece of evidence is that despite the *–a* suffix (definite article) appearing on previously feminine nouns, all associated words are inflected as masculine in this dialect. Thus, the pattern is *en bok* "a.M book," but *boka* "the book" (with the definite suffix for feminines). All adjectives and possessives are masculine, with the exception of certain instances of postnominal possessives. Together, these facts indicate that the gender of these nouns is M and that the suffix is indicating something that is not gender. Lødrup (2011), following Enger (2004), argues that the suffix expresses declension class, the inflection that is used for definite forms. As Alexiadou (2004, p. 25) points out, "[…] inflection class […] is never relevant for the purposes of agreement. It merely groups nouns into classes, which do not determine any further properties." In essence, then, the distinction between gender markers and declension class markers is based on different properties: The latter is always a bound morpheme and appears on the noun itself, whereas the former do not appear on the noun. Following Corbett and Fedden (2015), it could be argued that in systems where gender markers and declension class markers align, we have a *canonical* gender system, whereas the Oslo dialect exhibits a non-canonical gender system, where the definiteness suffix does not encode gender.

Gender is also marked on possessives, which may be either pre- or post-nominal. Note that the noun is marked for definiteness when the possessor appears after the noun. In contrast, the definite suffix is impossible if the possessor is prenominal. According to Anderssen and Westergaard (2012), who have investigated both the NoTa corpus of adult speech (Oslo)[3] as well as a corpus of child-directed speech recorded in Tromsø (Anderssen, 2006), the frequency of the postnominal possessor construction is much higher than the prenominal one (attested approximately 75%). The proportion of the postnominal possessor construction has been found to be even higher in American Norwegian heritage language, as the majority of the speakers investigated ($N = 34$) produce virtually only this word order (Westergaard and Anderssen, 2015). This is relevant for our investigation of gender, as it has been argued

---

[3]NoTa (Norsk talespråkskorpus—Oslodelen [Norwegian spoken corpus, the Oslo part]), *The Text Lab, Department of Linguistics and Scandinavian Studies*. University of Oslo. Available online at: http://www.tekstlab.uio.no/nota/oslo/index.html

that the possessor is not an exponent of gender when it is placed postnominally (cf. Lødrup, 2011). This means that it could be treated like a declension class marker just like the definite suffix, and as just mentioned, the postnominal possessive also retains the feminine form much more than the prenominal one. We return to this in the Section Our study: Participants, Hypotheses and Methodology.

Finally, we should note that Norwegian exhibits a phenomenon called double definiteness, requiring that definiteness be marked twice in certain contexts, notably in demonstratives and in modified noun phrases. This means that definiteness is marked both on a pre-nominal determiner and on the suffix. While double definiteness adds complexity to the Norwegian noun phrase, it is also worth noting that in case of the prenominal determiner, there is again syncretism between M and F forms (cf. **Table 1**).

GRAMMATICAL GENDER IN ACQUISITION AND ATTRITION

The Acquisition of Gender

Grammatical gender is a complex linguistic phenomenon. A child or a second language learner acquiring a language with gender thus often has to internalize a range of different cues that contribute to determining the gender of a given noun. For the acquisition of grammatical gender in Norwegian, the lack of transparency of gender assignment has been shown to be a major challenge. While gender is typically acquired around the age of three in languages with a transparent gender system, such as Russian (e.g., Gvozdev, 1961) or many Romance languages (e.g., Eichler et al., 2012, on various bilingual Romance-German combinations), gender has been shown to be in place relatively late in Norwegian. Based on corpora of two monolingual and two bilingual (Norwegian-English) children (age approximately 2–3), Rodina and Westergaard (2013) found considerable overgeneralization of masculine forms (by far the most frequent forms in the input) to both feminine and neuter nouns (63 and 71% respectively). In a more recent experimental study of somewhat older children and adults, Rodina and Westergaard (2015a) find that neuter gender is not in place (at 90% accuracy; cf. Brown, 1973) until the age of approximately 7. It is also shown that the feminine is even more vulnerable among the older children. Rodina and Westergaard argue that this latter finding is due to an ongoing change in the dialect (Tromsø) from a three-gender system to a two-gender system, common and neuter. In both studies, they also show that, while proper gender forms such as the indefinite article are late acquired, the corresponding declension class markers (e.g., the definite suffix) are target-consistently in place from early on. In fact, the acquisition pattern for indefinite and definite forms are the mirror image of one another at an early stage, with non-target-consistent production around 90% for the former category and only about 10% for the latter. This means that young children typically produce the masculine form of the indefinite article with nouns of all three genders (e.g., *en.M hest*(M) "a horse," *en.M seng*(F) "a bed," *en.M hus*(N) "a house," cf. **Table 1**), while the

definite suffix is target-consistent (*hesten* "the horse," *senga* "the bed," *huset* "the house"). Results confirming this pattern are also attested in an experimental study of bilingual Norwegian-Russian children (Rodina and Westergaard, 2015b). These findings show that learners do not create an immediate link between the definite suffix and the agreement forms, indicating that the two belong to different systems and thus support the distinction between gender and declension class in Lødrup (2011).

Gender in Heritage Language Situations

Over the past 20 years, there has been an increasing focus on the language of heritage speakers. We adopt the following definition of a heritage language: "A language qualifies as a heritage language if it is a language spoken at home or otherwise readily available to young children, and crucially this language is not a dominant language of the larger (national) society (Rothman, 2009, p. 156; see also e.g., Rothman, 2007; Polinsky, 2008; Benmamoun et al., 2013). One characteristic of heritage grammars is that they may be different from that of speakers acquiring the same language as a majority language due to incomplete acquisition (e.g., Polinsky, 1997, 2006; Montrul, 2002, 2008; Sorace, 2004; Tsimpli et al., 2004) or attrition (e.g., Pascual y Cabo and Rothman, 2012; Putnam and Sánchez, 2013). That means that a heritage language grammar may represent a change compared to the grammar of the previous generation as well as the relevant non-heritage variety.

The baseline language for a heritage speaker is the language of exposure during childhood. This means that a heritage speaker of Russian in the US should not strictly speaking be compared to a speaker of Russian in Russia. This makes studying heritage languages quite challenging, given that it is often difficult to establish the relevant properties of the primary linguistic data that the learners have been exposed to. Due to this lack of data across generations, a comparison is often made between the heritage language grammar and the non-heritage variety—with the caveat that the latter does not necessarily represent the input to the generation of heritage speakers studied. This is what we have had to do in the current study. Heritage speakers also differ from non-heritage speakers of the same language with respect to the amount of variation attested in their production; while some speakers have a fairly stable grammar, others display a more variable grammar, not applying rules consistently (see Montrul, 2008 for discussion).

It is well known that for heritage speakers, the amount of input and use of the language during childhood varies (see Montrul et al., 2008, among many others). Given the complexity of gender, it is to be expected that heritage speakers face difficulties with this part of the grammar. This has been investigated for Russian heritage language in the US by Polinsky (2008). Like Norwegian, Russian has three genders: masculine, feminine, and neuter; see Corbett (1991, pp. 34–43) and Comrie et al. (1996, pp. 104–117) for further details and references. According to Corbett (1991, p. 78) the distribution of the three genders is M 46%, F 41%, and N 13%. Gender agreement is marked on adjectives, participles, demonstratives, possessive pronouns, past tense verbs and some numerals, and gender assignment is relatively transparent in that M nouns typically end in a consonant, F nouns in –*a*,

and N nouns in –*o*. There are also some classes of nouns with non-transparent gender assignment.

Given somewhat reduced input, heritage speakers are typically exposed to fewer cues for gender assignment than children learning non-heritage Russian. Polinsky (1997, 2006) shows that less proficient American Russian speakers do not fully master the complex system of declension classes. In Polinsky (2008, p. 55), she demonstrates that two new gender systems have developed among the heritage speakers, both somewhat different from that of the non-heritage variety: (1) a three-gender system used by the more proficient speakers, differing from the non-heritage variety in that opaque N nouns ending in an unstressed –*o* are produced with F gender (i.e., they are pronounced with a schwa and therefore confused with the feminine ending –*a*), and (2) a two-gender system produced by the less proficient speakers where all N nouns have migrated to F. It is speculated that the latter speakers do not master the complex system of declensional case endings, and in the absence of this knowledge, they are relying on a purely phonological cue, i.e., whether the noun in its base form (Nominative singular) ends in a consonant or a vowel. The two systems are described in (1)–(2).

(1) More proficient speakers: Three-gender system

    a. nouns ending in a consonant are M

    b. nouns ending in a stressed –*o* are N

    c. all other nouns are F (i.e., including nouns ending in an unstressed –*o*, which are N in non-heritage Russian)

(2) Less proficient speakers: Two-gender system

    a. nouns ending in a consonant are M

    b. nouns ending in a vowel are F

In a recent study of Norwegian-Russian bilingual children growing up in Norway (age 4–8), Rodina and Westergaard (2015b) find an even more reduced gender system in some of the children. The amount of input is argued to be crucial: While children with two Russian-speaking parents are virtually identical to monolingual children growing up in Russia, the bilinguals with the least amount of input (only one Russian-speaking parent who does not use Russian consistently with the children) have considerable problems with gender, not just the opaque nouns, but also the transparent ones. In fact, some of these children produce almost exclusively masculine forms, overgeneralizing them to feminine nouns 77% and to neuters as much as 94%, which means that they do not seem to have any gender distinctions at all. Since these children are only up to 8 years of age, follow-up studies are necessary in order to find out whether they will eventually converge on the target, or whether they are developing a Russian heritage variety without gender.

Gender and Diachronic Change

It is well known that M and F genders have collapsed into common gender (C) in many Germanic languages and dialects. This change has taken place e.g., in Dutch, Danish, and the Bergen dialect of Norwegian (Jahr, 1998; Nesse, 2002; Trudgill, 2013). Furthermore, Conzett et al. (2011) have attested a similar

change in certain dialects in North Norway (Kåfjord and Nordreisa). This region has had extensive language contact with Saami and Kven/Finnish, languages which do not have grammatical gender. This language contact is argued to have caused a reduction of the gender system of the Norwegian spoken in this area from three to two (C and N). At the same time the declension system is intact. This means that while the feminine indefinite article *ei* "a.F" is virtually non-existent in the data, the corresponding definite suffix still has the *–a* ending typical of F nouns. This is illustrated in (3).

(3) a. en bok   - boka
     a.C book.C   - book.F.DEF

This pattern is identical to what Lødrup (2011) found for Oslo speech (cf. the Section Gender and the Norwegian Noun Phrase above). The cause of the change in Oslo is generally argued to be sociolinguistic: The *Bokmål* written standard allows the use of only two genders, and a spoken version of this variety enjoys a high social prestige in certain speaker groups. Thus, the three-gender system of the traditional dialects has gradually become associated with something rural and old-fashioned. The pattern attested means that a reduced gender system has developed in both areas (common and neuter), but at the same time a more complex declension system, in that the new common gender has two declension classes in the definite form, i.e., *en bil–bilen* "a car–the car" and *en bok–boka* "a book–the book."

Even more recent research is providing us with data on a real-time case of language change. Based on an experimental study, Rodina and Westergaard (2015a) demonstrate that F gender is rapidly disappearing from the speech of children and young adults in Tromsø: The F indefinite article is replaced by M, yielding common gender, but as in Oslo and Kåfjord/Nordreisa, the definite suffix is still preserved in its F form. Note that this pattern is also identical to what has been attested in early Norwegian child language (cf. the Section The Acquisition of Gender). While Rodina and Westergaard (2015a) also assume that the cause of this change is sociolinguistic, they argue that the *nature* of the change is due to acquisition: While the N forms are saliently different from the other two genders, there is considerable syncretism between M and F (e.g., adjectives and prenominal determiners), making it more difficult to distinguish the two in the acquisition process (cf. **Table 1**). Furthermore, while the real gender forms are very late acquired (around age 5–7), the declensional suffixes are target-consistently in place very early (around age 2), cf. Anderssen (2006) and Rodina and Westergaard (2013). Thus, the late acquired forms are the ones that are vulnerable to change.

The three studies briefly presented here demonstrate that F gender is disappearing or already lost from several Norwegian dialects. We would thus expect that F gender should be vulnerable in an acquisition context where there is somewhat reduced input, e.g., in a heritage language situation. In the following sections, we present our study of gender in American Norwegian.

OUR STUDY: PARTICIPANTS, HYPOTHESES, AND METHODOLOGY

Norwegian Heritage Language in America

According to Johannessen and Salmons (2012, p. 10), Norwegian immigration started in 1825, when the first Norwegians arrived in New York. By 1930, as many as 810,000 people had arrived in the US and an additional 40,000 in Canada. In the US, they settled mostly in the Midwest, predominantly in the Dakotas, Illinois, Iowa, Minnesota, and Wisconsin. The Norwegians built churches and schools and also had their own newspapers, *Decorah-Posten* and *Nordisk Tidende*. According to Johannessen and Salmons (2012, p. 6) 55,465 people reported Norwegian as their home language in the 2000 US Census. However, most of the current heritage speakers are above 70 years of age. American Norwegian as a heritage language can thus be said to be in its final stages (cf. Johannessen and Salmons, 2012).

American Norwegian was first documented and studied by Haugen (1953), based on fieldwork in the late 1930s and 1940s and subsequently, this heritage language was studied by Hjelde (1992, 1996). More recently, extensive fieldwork has been conducted in connection with the NorAmDiaSyn project, and data have been collected from a number of 2nd to 4th generation immigrants who learned Norwegian as their L1 from parents and grandparents. According to Haugen (1953, p. 340), the first immigrants were from the west coast of Norway, but around 1850, large numbers came from rural Eastern parts of Norway (Johannessen and Salmons, 2015, p. 10). It is mainly these Eastern varieties that are spoken today: Johannessen and Salmons (2015) remark that in 2010 it was difficult to find speakers of western dialects. For most of the immigrants, there was little or no support for Norwegian language in the community. Consequently, these speakers have generally been bilingual since the age of 5–6, and they have been dominant in English since this time. The background information offered about the corpus participants is relatively sparse: Year of birth, language of schooling and confirmation, literacy in Norwegian, number of visits to Norway as well as other contact with the country. In addition, we know which generation immigrant they report to be, and for some of them, the year their family arrived in the US. There is no information about the amount of use of Norwegian in adulthood. The language of schooling is English for all of them (except two informants for which this information is missing), and the large majority (43/50) had their confirmation in English. Contact with Norway varies between "some" and "often," and many have never visited the country. Typically, these heritage speakers have never had any instruction in Norwegian, and most of them have no literacy skills in the language.

The majority of the participants are between 70 and 100 years old today, and as they have not passed on the language to the next generation, they do not have many people to communicate with in Norwegian. Thus, most of these heritage speakers hardly ever use Norwegian any more, and at the time of the CANS recordings, many of the participants had not uttered a word of Norwegian for years, one participant for as long as 50 years. The initial impression of their Norwegian proficiency is that it is quite rusty, but once these speakers warm up, many properties

of the language turn out to be intact (Johannessen and Laake, 2015). Given the language profile of these learners (monolingual Norwegian speakers until school age, predominantly English dominant in adult life, and hardly using Norwegian at all in old age) it is possible that any discrepancies between their language and the non-heritage variety should be due to attrition rather than incomplete acquisition.

So far, data from 50 informants have been transcribed and now make up the Corpus of American Norwegian Speech (CANS) (Johannessen, 2015). This corpus consists of speech data collected through interviews (by an investigator from Norway) and conversations among pairs of heritage speakers. Each recording lasts approximately a half hour to an hour, meaning that there is relatively sparse data per informant.

Hypotheses and Predictions

Based on the properties of the gender system of Norwegian and previous research on gender in acquisition and change, we formulate the following hypotheses and predictions for American Norwegian:

(4) Hypotheses

 A. Gender is vulnerable in American Norwegian
 B. Gender forms and declensional suffixes behave differently
 C. F is more vulnerable than N due to syncretism with M

(5) Predictions

 A. Speakers will overgeneralize M gender forms
 B. Declensional suffixes will be retained
 C. F will be affected first; i.e., (some speakers of) American Norwegian will have a two-gender system (common and neuter)

We expect gender to be vulnerable in a situation with reduced input such as Norwegian heritage language, especially given the non-transparency of the gender system and the relatively late acquisition attested by Rodina and Westergaard (2015a). We also expect to see a difference between forms that express gender proper (i.e., agreement) and the declensional endings, which has been attested in previous research on both acquisition and change (e.g., Lødrup, 2011; Rodina and Westergaard, 2013). Finally, as in Russian heritage language and in many Germanic varieties, we may also see reductions in the gender system, either from a three- to a two-gender system (common and neuter) or to a system where gender breaks down completely.

Methodology

We have used CANS to probe the usage of gender in American Norwegian. We have generally excluded English loan words appearing with gender marking (see Flom, 1926; Hjelde, 1996; Nygård and Åfarli, 2013; Alexiadou et al., 2015 on this issue)[4].

---

[4]It is not always easy to distinguish loan words from English words that have become an integrated part of American Norwegian speech, e.g., *farmer* or *field*. We have used the following criterion in our selection: All words that currently exist in English and which are pronounced with a clear American pronunciation have been discarded in this paper.

Our main focus here is on gender assignment, and we have therefore also disregarded agreement between different gender forms within the nominal phrase. We have searched CANS for the following forms:

(6) a. the indefinite article followed by a noun (occasionally with an intervening adjective)
 b. possessives
 c. definite forms

We have also compared the data from the CANS corpus to a sample of the Nordic Dialect Corpus (Johannessen et al., 2009). This allows us to compare the gender system of American Norwegian to that of contemporary Norwegian. We would like to emphasize that we obviously do not assume that the heritage speakers recorded in the CANS corpus were exposed to a variety of Norwegian that is identical to the non-heritage variety spoken today. But we are interested in investigating possible changes in the heritage variety, possibly across several generations, and these are the data we have available to make the comparison. We have used the part of the Nordic Dialect Corpus which covers the dialects spoken in the Eastern part of Norway (excluding the capital, Oslo), the area from which most of the ancestors of the heritage speakers originate. The Nordic Dialect Corpus consists of structured conversations between speakers of the same dialect and as such, the two corpora are comparable with respect to the recording situations. In the Nordic Dialect Corpus, speakers are classified as either "old" (over 50) or "young" (under 30), where most of the informants in the two groups are in their 60s and 20s respectively. The corpus was recorded between 2008 and 2011.

Both corpora have been transcribed into a dialect version and a standardized *Bokmål* transcription. The corpora are tagged, and the transcriptions are directly linked to the recordings. In CANS, we found that in several cases, the *Bokmål* transcription had standardized the gender according to the *Bokmål* official dictionary, even when the informants actually used a different gender. Thus, we have had to check the recordings carefully in order to be sure that we had reliable transcriptions. We generally did not find errors in the dialect version (corresponding to the pronunciation), which made us trust that this transcription is sufficiently correct for our present purposes. Furthermore, there are some instances where the F indefinite article has been transcribed simply as /e/. We have listened to all of these and in all cases the informants seem to be saying the feminine form /ei/. They have therefore been counted as occurrences of the F indefinite article.

Compound nouns (e.g., *skolehus* "school house") have been counted separately. In Norwegian, the right-hand part of the compound is always the head noun and thus determines the gender. For several of the compound words in the corpus, the right-hand noun also occurs independently (e.g., *hus* "house"). Instances where the noun was not uttered completely were disregarded. In cases where speakers correct themselves as in (7a), we only counted the latter form. Examples have also been counted if they occur in what would be considered an ungrammatical or unidiomatic structure in Norwegian, e.g.,

(7b), which is presumably a direct translation of an English expression.

(7) a. *ei # en familie*   (flom_MN_02gm)
       a.F # a.M family.M

   b. *vi hadde en god tid*   (portland_ND_01gm)
       we had a.M good time.F
       Target form (intended meaning): *vi hadde det morsomt*
       [lit.: we had it fun]

With these methodological considerations in mind, let us move on to the results of our study.

RESULTS

Gender Marking on the Indefinite Article—Overall Results

Our search in CANS first of all revealed that all three gender forms are attested in the data. Examples illustrating the use of the three indefinite articles *en, ei,* and *et* (M, F, and N) are provided in (8)–(10). In these examples, the gender marking is entirely in line with what we would expect in present-day non-heritage Norwegian. It is also worth noticing that although there is some language mixing between English and Norwegian here, the sentences are predominantly Norwegian in structure and lexicon.

(8) *vi kjøpte **en butikk***   (blair_WI_04gk)
    we bought a.M store(M)
    "we bought a store"

(9) *og **ei uke**   sia så h- visita vi parken   i Blair*
    and a.F week(F) ago so visited   we park.DEF in Blair
    her   (blair_WI_01gm)
    here
    "a week ago we visited the park in Blair here"

(10) *we got har bare **et tre***   (coon_valley_WI_04gm)
     we got have only a.N tree(N)
     "we only got one tree"

In a study of the *Nynorsk* dictionary (Hovdenak et al., 1998), which is the written norm that is closest to the contemporary dialects, Trosterud (2001) has found that out of the 31,500 nouns listed there, 52% are M, 32% are F, and 16% are N. These numbers are somewhat different from the distribution in the spoken language. Rodina and Westergaard (2015a) have investigated proportions of the indefinite article in a corpus of child and child-directed speech recorded in the mid-90s (Anderssen, 2006) and found that M forms are even more frequent in the input than in the dictionary, 62.6%, while the F and N forms are more or less equally represented, 18.9 and 18.5% respectively (N = 2980). We have investigated the occurrences of the three indefinite articles in the Nordic Dialect Corpus, and we find that the distribution in the data of the "old" speakers is virtually identical to Rodina and Westergaard's (2015a) findings, see **Table 2**. In the data of the "young" speakers, on the other hand, the F indefinite article is only attested 5.4%, while the proportion of M forms has increased to 74.9%. We believe that it is likely that these numbers reflect

TABLE 2 | Token distributions of the three indefinite articles *en* (M), *ei* (F) and *et* (N), in CANS and in Eastern Norwegian dialects (Nordic Dialect Corpus).

Gender	CANS (N = 50)	NorDiaCorp (old, N = 127)	NorDiaCorp (young, N = 66)
M	76.3% (753)	64.8% (1833)	74.9% (909)
F	16.9% (165)	18.2% (514)	5.4% (66)
N	6.9% (67)	17.0% (481)	19.7% (239)

an ongoing change involving the loss of F forms also in these dialects, just like in Oslo and Tromsø (cf. the Section Gender and Diachronic Change). A careful study of the Nordic Dialect Corpus in order to confirm (or disconfirm) this hypothesis has to be left for future research.

In **Table 2**, we have also provided the relevant counts from the CANS corpus. Overall, the figures for the heritage speakers indicate that gender is relatively stable in American Norwegian, as they are quite similar to the older speakers in the Nordic Dialect Corpus, except for a lack of neuter forms. However, a closer look reveals that the heritage speakers are overgeneralizing the M gender forms quite substantially to both F and N nouns. We now turn to a discussion of these discrepancies between the CANS corpus and forms found in present-day spoken Norwegian.

Overgeneralization—Indefinite Articles

Although all gender forms are represented in the corpus, and gender thus appears to be relatively stable, there are several cases of what we will refer to as non-target-consistent forms, i.e., forms that are different from what would be expected in non-heritage Norwegian. When determining the gender of nouns in non-heritage Norwegian, we have used the *Nynorsk* Dictionary with some adjustments for differences between the dictionary and the gender typically found in Eastern Norwegian dialects[5]. In this section, we consider nouns with the indefinite article, either by itself or together with an adjective. We first consider all noun occurrences (tokens) and then the number of different nouns (types) appearing in the corpus.

In the corpus, we find 236 occurrences that are F nouns. As many as 39.0% (92/236) of these appear with M gender; see (11)–(13).

(11) *og   om   in in **en uke**   da   så # kom han til byen igjen*
     and about in in a.M week.F then so came he to city again
                              (rushford_MN_01gm)

(12) *og # tre brødre og   r- en s- # **en søster***
     and   three brothers and   a.M a.M sister.F
                              (blair_WI_04gk)

---

[5]We are grateful to Jan Terje Faarlund for valuable help and discussions concerning this issue.

(13) *ja # em # har du # har du **en ku** enda?*
     yes em   have you have you a cow still?

                                 (coon_valley_WI_01gk)

We should note that there is considerable variation between M and F forms used with some F nouns in the corpus. For example, *datter* "daughter" occurs both with F and M indefinite articles. Speakers appear to be consistent and typically do not alternate. However, given the sparse data in CANS, we very often find that a speaker only produces one or two instances of the same noun. For this reason, we cannot address the question of speaker consistency.

Turning to the neuter, we find 164 nouns which are N according to the *Nynorsk* dictionary and our Eastern Norwegian adjustments. Of these, as many as 48.8% (80/164) appear with the M indefinite article. Examples are provided in (14)–(16).

(14) *# bestemor # var født # hun var på **en fjell** ##*
     grandma   was born she   was on a.M mountain.N

                                   (chicago_IL_01gk)

(15) *fire år # og en **en år***   at to the university
     four years and a.M a.M year.F at to the university

                                 (wanamingo_MN_04gk)

(16) *fikk jeg **en pass***   (coon_valley_WI_02gm)
     got I   a.M passport.N

There are also occasional N nouns appearing with F gender forms, 10.4% (17/164); see the examples in (17)–(19). Considering the current trend in Norway with F gender in the process of disappearing, it is rather surprising that there is overuse of feminine forms.

(17) *det var ei menneske*   (westby_WI_05gm)
     it   was a.F human.being.N

(18) *han var her det var ei bryllup*
     he was here it   was a.F wedding.N

                                   (harmony_MN_02gk)

(19) *jeg tror ikke jeg sa ei eneste norsk   ord*
     I   think not I   said a.F single Norwegian word.N

                                   (harmony_MN_04gm)

Finally, we found four examples of non-target-consistent gender on M nouns, in all cases produced with the F indefinite article. This amounts to only 0.7% (4/576).

We now take a closer look at the number of actual nouns involved (types). Due to the very low number of non-target-consistent M nouns, we only consider F and N. The list in (20) provides all F nouns that occur with the target-consistent indefinite article (altogether 51 nouns), where the ones in bold are sometimes produced with M (10 nouns). In (21) we find 21 F nouns that *always* appear with M gender in the corpus. In

total, there are 72 different F nouns, of which 31 are either always or sometimes produced with M gender forms. This means that overgeneralization of types is 43.1% (31/72), which is similar to the frequency of noun tokens reported above, 39.0%.

(20) **F=F:** *stund* "time," *søster* "sister" *kanne* "mug," *trå* "yearning," **side** "side," **kjerring** "hag," **seng** "bed," **uke** "week," *jente* "girl," *lefse* "lefse," **kiste** "coffin," *mølle* "mill," *øks* "ax," **tid** "time," *mjølking* "milking," **ku** "cow," *kvige* "heifer," *grøft* "trench," *brødpanne* "bread pan," **bok** "book," *trinse* "caster," *mil* "mile," *høstnatt* "fall night," *datter* "daughter," *dame* "lady," *bjelle* "bell," *tobakksseng* "tobacco bed," *ei [female name removed]* "a female name," *bestemor* "grandmother," *hytte* "hut," *frilledatter* "daughter of a mistress," *gryte* "pot," *aure* "trout," *liste* "list," *skrøne* "tall tale," *rumpe* "butt," *stikke* "peg," *pakke* "package," *pike* "girl," *mor* "mother," *trønderskrøne* "tall tale from Trøndelag," *dør* "door," *plattform* "platform," *himmelseng* "four-poster bed," **kirke** "church," *tante* "aunt," *hand* "hand," *matte* "mat," *lue* "cap," *bøtte* "bucket," **datter** "daughter" (41 + 10 = 51)

(21) **F->M:** *blanding* "mixture," *mil* "mile," *flaske* "bottle," *tale* "speech," *stund* "while," *gruppe* "group," *ordbok* "dictionary," *hast* "haste," *rotte* "rat," *vogn* "wagon," *avis* "newspaper," *pipe* "pipe," *elv* "river," *stripe* "stripe," *kagge* "keg," *purke* "sow," *slekt* "family," *øy* "island," *dialekt* "dialect," *klasse* "class," *lærerinne* "female teacher" (21)

Considering N nouns, (22) lists all the ones that occur with the target-consistent indefinite article (altogether 23 nouns). Nouns in bold also appear with M indefinite article (11 nouns), while nouns which are underlined also appear with F (8 nouns). In (23) we find N nouns which only appear with F indefinite article and in (24) N nouns that consistently appear with M indefinite article.

(22) **N=N:** **hotell** "hotel," <u>**par**</u> "pair/couple," *år* "year," *fat* "plate," <u>**brev**</u> "letter," *lass* "load," *hus* "house," <u>lag</u> "layer," <u>hull</u> "hole," <u>**skolehus**</u> "school," **bilde** "picture," *sted* "place," <u>**fjell**</u> "mountain," <u>blad</u> "magazine," <u>ord</u> "word," <u>rom</u> "room," *leven* "noise," **stykke** "piece," *slag* "blow," **navn** "name," *minutt* "minute," **liv** "life," *problem* "problem" (12 + 11 = 23)

(23) **N->F:** *menneske* "human being," *hjem* "home," *bryllup* "wedding," **barnebarn** "grandchild," *papir* "paper" (5)

(24) **N->M:** *barnetog* "children's parade," *farmeår* "farm year," *program* "program," *pass* "passport," *tømmerhus* "log cabin," *tog* "train," *arbeid* "work," *patent* "patent," *dusin* "dozen," *bord* "table," *band* "band," *lys* "light," *oppstuss* "fuss," *eiketre* "oak," *utvandrermuseum* "emigration museum," *kort* "card," *mål* "measure," *måltid* "meal," *kupp* "bargain," *selvfirma* "independent company," *orkester* "orchestra" (21)

The total number of different N nouns is 49. As many as 34 of them (always or sometimes) appear with an M indefinite article (69.4%), while 13 (always or sometimes) appear with F gender (26.5%). This means that N nouns are quite unstable in the production of these heritage speakers.

TABLE 3 | Summary of noun tokens and noun types appearing with a non-target-consistent indefinite article.

Direction	Total number of examples (tokens)	Number of different nouns (types)
F → M	39.0% (92/236)	43.1% (31/72)
N → M	48.8% (80/164)	69.4% (34/49)
N → F	10.4% (17/164)	26.5% (13/49)

Table 3 summarizes our findings, considering both the total number of noun occurrences (tokens) in the data as well as the number of different nouns (types).

Gender vs. Inflection Class

As we have seen, many of the F and N nouns in the corpus (always or sometimes) occur with an M indefinite article (31/72 and 34/49 respectively), shown in (25) and (27). However, when we consider the definite suffixes on these same nouns, they are usually the feminine –*a* and neuter –*et* forms, not the masculine –*en*. This is shown in (26) and (28), where the numbers in parentheses indicate occurrences. In fact, for the neuter nouns, the masculine declensional suffix is unattested (cf. Johannessen and Larsson, 2015).

(25) *en datter* "a daughter," *en tid* "a time," *en kirke* "a church," *en uke* "a week"

(26) *dattera* (24) - *datteren* (0), *tida* (206) - *tiden* (13), *kirka* (80)- *kirken* (3), *uka* (14) - *uken* (0)

(27) *en år* 'a year', *en tog* "a train," *en hus* "a house," *en lys* "a light"

(28) *året* (31) – *åren* (0), *toget* (9) – *togen* (0), *huset* (60) – *husen* (0), *lyset* (3) – *lysen* (0)

This mirrors findings from other studies, showing that when the feminine gender is lost, the definite suffix is retained (e.g., Lødrup, 2011; Rodina and Westergaard, 2015a). This demonstrates that the affixal definite article clearly behaves differently from the free gender morphemes that agree with the noun, e.g., the indefinite article, not only in contexts of acquisition and change, as attested in previous research, but also in heritage language.

Related to this is the result of our search for possessives in the corpus. Recall from the Section Gender and the Norwegian Noun Phrase that possessives in Norwegian may appear both in prenominal and postnominal position, and that Westergaard and Anderssen (2015) reported that in Norwegian heritage language, the postnominal construction is the preferred one. First of all, our findings show that the possessives used in the corpus are mainly high frequency kinship terms (more than 90%) of the type illustrated in (29)–(30); thus, they may be rote-learned or memorized and not necessarily be the result of a productive system. We also find that numbers are very low for all possessives except the first person singular, and this is therefore the only result that is reported here (**Table 4**).

(29) a. *mor   mi* (44)
        mother my
        "my mother"

TABLE 4 | Distribution of gender marking for first person possessives in CANS (*N* = 50).

Gender form	Prenominal	Postnominal	Total
M (*min* "my")	87/96 (90.6%)	251/414 (60.6%)	338/510 (66.3%)
F (*mi* "my")	0 (–)	126/414 (30.4%)	126/510 (24.7%)
N (*mitt* "my")	9/96 (9.4%)	37/414 (8.9%)	46/510 (9.0%)

      b. *søstera   mi* (10)
         sister.DEF my
      "my sister"

      c. *bestemora   mi* (4)
         grandmother.DEF my
      "my grandmother"

(30) a. *far   mi*n (102)
        father.DEF my
       "my father"

      b. *bror   min* (36)
         brother.DEF my
      "my brother"

      c. *mannen   min* (35)
         husband.DEF my
      "my husband"

Compared to the results in **Table 2**, where the proportion of F indefinite articles was only 16.9%, it is a bit surprising that the proportion of F forms is as high as 24.7%. However, as we mentioned above, the postnominal possessor has been argued to be a declension class marker and not an exponent of gender (Lødrup, 2011). In this table, we also see that the prenominal possessives behave differently from the postnominal ones, in that the feminine form is attested relatively frequently as a declension class marker (30.4%), and not at all in the gender form (in prenominal position). This difference becomes even clearer when we consider whether the gender forms have been used target-consistently: In **Table 5**, the feminine forms are always produced with M gender in prenominal position (the gender form) but they are generally retained when occurring postnominally, where we only find occasional non-target forms (both M and N). The fact that the F form is retained postnominally fits well with Lødrup's (2011) analysis that postnominal possessors behave like declension markers on a par with the affixal F definite endings. Turning to N nouns, we see that they also tend to migrate to M, somewhat more in prenominal than postnominal position (30.8 vs. 19.2%). In comparison, the masculine is virtually always produced with target-consistent gender agreement.

Individual Results

The individual production results of each of the 50 participants in the corpus are provided in the Appendix, for the indefinite article only, as this is the most frequent form produced. As expected, there is a very limited amount of data per informant, so that it is impossible to provide complete profiles of the gender system of each of them. Nevertheless, the participants have been

TABLE 5 | Distribution of genders for first person possessives in CANS (*N* = 50), target and non-target-consistent forms.

Gender	Prenominal		Postnominal	
	Target	Non-target	Target	Non-target
M	40/40	0	226/228	2 (to N, 0.9%)
F	0	43 (to M, 100%)	126/137	7 (to M, 5.1%)+4 (to N, 2.9%)
N	9/13	4 (to M, 30.8%)	21/26	5 (to M, 19.2%)
TOTAL	49	47	373	18

divided into four groups. In Group 1, there are four participants for which no conclusions can be drawn, as the production is too limited (one participant produces no indefinite forms at all and three participants only produce masculine forms—for masculine nouns). In Group 2, we find five participants who may possibly have an intact three-gender system, as they make no mistakes. However, each of them produces so few examples (11, 13, 9, 6, 6 respectively), and it is therefore possible that this is simply the result of sheer luck in the recording situation. Furthermore, only two of these five produce nouns in all three genders, while the remaining three only produce masculine and feminine nouns, not a single neuter. At the other end of the scale, there are nine informants who may not have gender at all (Group 3). These speakers produce masculine forms only, either for nouns belonging to two of the genders (four participants) or all three (five participants). The final group (Group 4) thus contains the majority of informants (32), who produce a mixture of forms. For these, target-consistency varies considerably, from participants making only one mistake (e.g., decorah_IA_01gm), who are thus similar to Group 2, to those who produce only one form that is not masculine (e.g., portland_ND_02gk) and are thus similar to Group 3. There is also variation with respect to which gender is more vulnerable, as some seem to have more problems with feminine nouns (e.g., webster_SD_02gm) and others with the neuter (e.g., coon_valley_WI_06gm), while others again have problems with both (e.g., stillwater_MN_01gm). Eight informants produce no feminine forms, which at first sight could indicate that they have a two-gender system consisting of common and neuter. However, two of them do not produce any feminine nouns at all, and all of them also make a considerable number of mistakes with the neuter. Thus, not a single informant displays a clear two-gender system where the neuter is intact and the feminine has merged with the masculine into common gender.

DISCUSSION

We now return to our hypotheses and predictions, repeated in (31)–(32) for expository convenience.

(31) Hypotheses

A. Gender is vulnerable in American Norwegian
B. Gender forms and declensional suffixes behave differently

C. F is more vulnerable than N due to syncretism with M

(32) Predictions

A. Speakers will overgeneralize M gender forms
B. Declensional suffixes will be retained
C. F will be affected first; i.e., (some speakers of) American Norwegian will have a two-gender system (common and neuter)

In the results section Gender Marking on the Indefinite Article–Overall Results, we saw that all the three genders are represented in the corpus, and the total numbers give the impression of a fairly stable system. However, when we considered the data in more detail (Section Overgeneralization—Indefinite Articles), we saw that there is considerable overgeneralization of M forms of the indefinite article to both F and N nouns (cf. **Table 3**). The substantial overgeneralization of M to F is unsurprising, given the findings from previous studies. However, in the present study there is clearly more overgeneralization affecting neuter than feminine nouns, both when we consider the overall number of occurrences (tokens, 48.8 vs. 39.0%) and the number of different nouns affected (types, 69.4 vs. 43.1%), cf. **Table 3**. In the prenominal possessives, we find that the feminines are produced with masculine forms 100% and the neuters approximately 31%. Based on these results, we conclude that gender is in fact vulnerable in American Norwegian, and thus that our Hypothesis A has been confirmed. Likewise, we can confirm Prediction A: Although there are a number of cases where neuter nouns migrate to the feminine (10.4% of the total number of neuters (tokens) and 26.5% of the number of different nouns (types), cf. **Table 3**), it is clear that the general pattern found for non-target-consistent forms is overgeneralization of the masculine.

Turning to Hypothesis and Prediction B, we saw in the Section Gender vs. Inflection Class that the definiteness suffix behaves very differently from the indefinite article. While feminine and neuter indefinite articles are frequently produced with masculine forms, the definite suffix is always target-consistent in the neuter and mostly also in the feminine. This means that our findings confirm previous research both from acquisition and change (cf. Sections The Acquisition of Gender and Gender and Diachronic Change), where the same distinction has been attested. As mentioned above, we consider the indefinite article to be an exponent of gender, whereas the affix is analyzed as a declension marker. The different behavior of these two elements also in this population of heritage speakers clearly shows that gender forms are much more prone to change than declension markers. The different behavior of the prenominal and postnominal possessives (at least for feminine nouns) also indicates that there is a distinction between the two that may be related to gender (cf. Lødrup, 2011).

It should be noted here that our claim that gender is vulnerable in Norwegian heritage language runs counter to the conclusion reached by Johannessen and Larsson (2015). Based on an investigation of a selection of the 50 speakers in CANS, they argue that grammatical gender is *not* affected by attrition. The main reason for the two different conclusions is that, unlike us, Johannessen and Larsson (2015) do consider the definite

suffix as a gender marker. And since the form of the suffix is generally retained, they consider this evidence that gender is intact. Furthermore, they find that complex noun phrases (determiner-adjective-noun) are much more prone to errors than simple ones (adjective-noun), with 18% (20/113) vs. 2% (1/58) target-deviant agreement. They argue that this shows that gender is unaffected by attrition, since it is target-consistent in simple noun phrases, and they account for the target-deviance in the complex ones as a result of processing difficulties. In our view, another explanation is also possible: Given that the number of noun types in the corpus is quite low and mainly consists of high-frequency nouns, we could argue that the simple noun phrases are more likely to be rote-learned and memorized as chunks than the more complex ones, which require a productive system of gender agreement. Since this is in the process of breaking down, the complex noun phrases display more errors.

We then turn to our final hypothesis and prediction (C) and the issue whether F gender is more vulnerable than N and whether we see changes or reductions in the gender system. As discussed above, this has been attested in Russian heritage language; both a reduction from a three- to a two-gender system (Polinsky, 2008) and possibly a breakdown of gender altogether (Rodina and Westergaard, 2015b). We also know that a reduction in the gender system has happened in many Germanic varieties and is currently taking place in certain Norwegian dialects (cf. the Section Gender and Diachronic Change), that is, a reduction from a three-gender system to a system with just two genders, common and neuter. As noted above, disappearance of also the neuter gender is not an unlikely scenario, given the non-transparency of the system and the late acquisition of this property of the Norwegian language. The gender system may be further weakened by the considerable lack of input and use in this heritage language situation. However, as shown in the previous section, we do *not* find any evidence of a two-gender system in the production of any these 50 speakers. Instead we see a general erosion across the whole gender system, with both feminine and neuter nouns migrating to the most frequent gender form, the masculine. In fact, the majority of the speakers ($N = 32$) behave in this way (Group 4). The end result of this will presumably be a complete breakdown of gender altogether; i.e., a system without gender distinctions. It is possible that this is already attested in the production of the nine speakers in Group 3, who produce only masculine forms.

We would like to speculate about the reasons for this development; i.e., (1) why is grammatical gender vulnerable in heritage language, (2) why are declension class suffixes stable, and (3) why do we not see evidence of a two-gender system the way we predicted? Our findings partly correspond to what has been found in acquisition and change, i.e., proper gender forms such as the indefinite article are late acquired and prone to change, while the declensional suffixes are early acquired and remarkably stable. But we do not find a two-gender system (common and neuter), which is attested in some children and which is also the result of changes that have taken place in certain varieties of Norwegian.

An obvious answer to the first question corresponds to the general account for the late acquisition of gender in Norwegian, viz. the non-transparency of gender assignment. A system where

gender has to be learned noun by noun is crucially dependent on a considerable amount of input. Unfortunately, we do not know much about the input to these speakers in childhood, but it is not inconceivable that it was somewhat limited. Given that gender has been found not to be fully in place until around age 6–7 (Rodina and Westergaard, 2015a), which is the time when these speakers experienced a language shift, it is possible that this property is the result of incomplete acquisition (e.g., Montrul, 2008). However, given the general profile of these heritage speakers mentioned above (monolingual Norwegian speakers until school age, English dominant in their adult lives, and hardly using Norwegian at all in old age), it is more likely that whatever discrepancies we find between their language and the non-heritage variety is due to attrition. This is further supported by the fact that there is considerable variation among these speakers. If this is the case, then we may speculate on a possible difference between incomplete acquisition and attrition with respect to gender: While the former process typically results in a systematic reduction in the gender system (e.g., from three to two genders), the latter affects an existing system in terms of erosion across the board. That is, incomplete acquisition is the cause of a system that is different from the non-heritage variety (and typically reduced), while the result of attrition is an unsystematic breakdown of the system, eventually leading to total loss of grammatical gender. Some support for our speculation may be found in Schmid's (2002) important work on German Jews in the United States, who had generally also experienced a severe reduction in the use of their L1 over an extended period of time: The occasional mistakes found in gender assignment in the data did not constitute any rule-based reduction in the gender system of their German[6].

We then turn to the second question, why declensional suffixes are stable in heritage language. The early acquisition of declensional suffixes is generally accounted for by their high frequency and the fact that they are prosodically favored by young children (Anderssen, 2006)[7]. They may also be initially learned as a unit together with the noun, even though they are not considered to be fully acquired until the relevant nouns also appear in appropriate contexts without the suffix. While prosody is unlikely to be a factor in heritage languages, the other two, frequency and chunking, may be responsible for the robustness of the definite forms. That is, highly frequent nouns (such as the ones typically used by our heritage speakers in the corpus) may be stored in memory as units together with the suffix, e.g., *hest**en*** "the horse," *seng**a*** "the bed," *hus**et*** "the house." For this reason, they are easily retrieved, while the indefinite forms must be computed as part of a productive process, e.g., *en hest* "a horse,"

---

[6]An important difference between Schmid's (2002) study and ours (pointed out by a reviewer) is that she finds very few non-target-like examples in her data, while there is evidence for considerable erosion in the data of the Norwegian heritage speakers. We would like to suggest that a possible reason for this could be that Schmid's (2002) subjects are first generation immigrants and thus had more robust input in their L1, while the attrition we see in our speakers could have accumulated over 3–4 generations. Furthermore, the German gender system could be said to be somewhat more transparent than the Norwegian one.

[7]Adding a definite suffix to monosyllabic nouns in Norwegian results in a trochaic structure (strong-weak), which is known to be favored by young children (e.g., Gerken, 1994).

ei seng "a bed," *et hus* "a house." In any case, our heritage data provide further evidence that the definite suffix does not have a gender feature. If this were the case, we would expect these speakers to make a direct link between this form and (other) gender forms: That is, knowing the definite form of a feminine or neuter noun (e.g., *boka* "the book" or *huset* "the house" should make it easy to produce the target-consistent indefinite forms *ei bok* "a book" and *et hus* "a house." But the data from these heritage speakers show that this is not the case. We therefore conclude that the evidence that we had from acquisition and change from previous studies is now supported by data from a new population.

Finally, we address the third question, why there is no systematic reduction from a three- to a two-gender system in the data of the heritage speakers. In several varieties of Norwegian that have undergone (or are undergoing) a change, the result has been the same: disappearance of the feminine and a development of a two-gender system with common and neuter gender. This has been argued to be partly due to sociolinguistic factors such as language contact or the prestige of the written form *Bokmål* and partly due to the syncretism between masculine and feminine, making it more difficult to distinguish the two in acquisition (e.g., Lødrup, 2011; Trudgill, 2013; Rodina and Westergaard, 2015a). Following up on our speculation above, we would like to suggest that all of these historical developments are due to incomplete acquisition. What we see in our data from the Norwegian heritage speakers, on the other hand, is the result of attrition. If this idea is on the right track, we might have a way to distinguish between the two processes: While incomplete acquisition typically results in a systematic difference between the heritage language and the non-heritage variety, attrition will result in general erosion and considerable variability[8].

CONCLUSION

In this paper, we have presented an investigation of grammatical gender in a corpus of heritage Norwegian spoken in America, the Corpus of American Norwegian Speech (CANS). The corpus consists of data from 50 speakers, whose linguistic profile is as

---

[8]A reviewer suggests that our findings could be the result of problems with lexical access in very old speakers rather than attrition. We agree that this could very well be the case—or at least an additional factor. This would predict that also Norwegians living in Norway would experience problems with gender assignment in their old age. Unfortunately, we know of no studies that have investigated this issue, and we therefore have to leave this suggestion to further research.

follows: Monolingual Norwegian until age 5–6, English dominant throughout life, and virtually no use of Norwegian in old age. Due to the non-transparency of gender assignment, we expected gender to be vulnerable in this situation of reduced input and use. Based on previous research from acquisition and change, we also expected declensional suffixes to be robust and feminine forms to be more vulnerable than neuter. That is, we expected to find evidence of a reduction in the system, from three genders (masculine, feminine, neuter) to two (common and neuter). Focusing on indefinite articles and possessives, we demonstrated that all three gender forms, masculine, feminine and neuter, are represented in the data. Nevertheless, there is considerable overgeneralization of masculine forms (the most frequent gender forms) in the production of the heritage speakers to both feminine and neuter nouns (as compared with gender in the relevant present-day Norwegian dialects). We also found a substantial difference between the indefinite article (an exponent of gender) and the definite suffixal article (which we consider a declension class marker): While the former is to a large extent affected by overgeneralization, the latter form is virtually always target-consistent. This confirms similar findings from previous research on both acquisition and change. However, we did not find any evidence of a two-gender system in the production of any of the speakers; instead there seems to be overgeneralization of masculine forms across the board. Assuming that the Norwegian of our participants is somewhat attrited, we speculate that this finding is due to a distinction between (incomplete) acquisition and attrition: While the former process typically results in a systematic difference between the heritage language and the non-heritage variety, attrition will lead to general erosion of the system and eventually complete loss of gender.

AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

REFERENCES

Alexiadou, A. (2004). "Inflection class, gender and DP-internal structure," in *Explorations in Nominal Inflection*, eds G. Müller, L. Gunkel and G. Zifonun (Berlin: Mouton de Gruyter), 21–50.

Alexiadou, A., Lohndal, T., Åfarli T. A., and Grimstad, M. B. (2015). "Language mixing: a distributed morphology approach," in *Proceedings of NELS,* Vol. 45, eds T. Bui and D. Özyildiz (Create Space), 25–38.

Anderssen, M. (2006). *The Acquisition of Compositional Definiteness in Norwegian.* Ph.D. dissertation, University of Tromsø.

Anderssen, M., and Westergaard, M. (2012). Tospråklighet og ordstilling i norske possessivkonstruksjoner. [Bilingualism and word order in Norwegian possessive constructions]. *Norsk Lingvistisk Tidsskrift [Nor. J. Linguist.]* 30, 170–197.

Benmamoun, E., Montrul, S., and Polinsky, M. (2013). Heritage languages and their speakers: opportunities and challenges for linguistics. *Theor. Linguist.* 39, 129–181. doi: 10.1515/tl-20 13-0009

Brown, R. (1973). *A first Language: The Early Stages.* Harvard, MA: Harvard University Press.

Comrie, B., Stone, G., and Polinsky, M. (1996). *The Russian Language in the Twentieth Century*. Oxford: Clarendon Press.

Conzett, P., Johansen, Å. M., and Sollid, H. (2011). Genus og substantivbøying i nordnorske språkkontaktområder. [Grammatical gender and declension in language contact areas in North Norway]. *Nordand Tidskrift for andrespråksforskning [Nordand J. Sec. Lang. Res.]* 6, 35–71.

Corbett, G. G. (1991). *Gender*. Cambridge University Press.

Corbett, G. G., and Fedden, S. (2015). Canonical gender. *J. Linguist*. doi: 10.1017/S0022226715000195. [Epub ahead of print].

Dahl, Ö. (2000). "Elementary gender distinctions," in *Gender in Grammar and Cognition II: Manifestations of Gender*, eds B. Unterbeck, M. Rissanen, T. Nevalainen, and M. Saari (Berlin: Mouton de Gruyter), 577–593.

Delsing, L.-O. (1993). *The Internal Structure of Noun Phrases in the Scandinavian Languages*. Ph.D. dissertation, University of Lund.

Eichler, N., Jansen, V., and Müller, N. (2012). Gender acquisition in bilingual children: French-German, Italian-German, Spanish-German and Italian-French. *Int. J. Bilingual*. 17, 550–572. doi: 10.1177/13670069114 35719

Enger, H.-O. (2004). On the relation between gender and declension: A diachronic perspective from Norwegian. *Stud. Lang*. 28, 51–82. doi: 10.1075/sl.28.1. 03eng

Enger, H.-O., and Corbett, G. (2012). Definiteness, gender, and hybrids: evidence from Norwegian Dialects. *J. German. Linguist*. 24, 287–324. doi: 10.1017/S1470542712000098

Faarlund, J. T., Lie, S., and Vannebo, K. I. (1997). *Norsk referansegrammatikk. [A Reference Grammar of Norwegian]*. Oslo: Universitetsforlaget.

Flom, G. T. (1926). English loanwords in American Norwegian, as spoken in the Koshomong settlement, Wisconsin. *Am. Speech* 1, 541–558.

Gerken, L. (1994). Young children's representation of prosodic phonology: evidence from English-speakers' weak syllable productions. *J. Mem. Lang*. 33, 19–38.

Gvozdev, A. N. (1961). *Formirovanie u Rebenka Grammatièeskogo Stroja Russkogo Jazyka. [Language development of a Russian child]*. Moscow: APN RSFSR.

Haugen, E. (1953). *The Norwegian Language in America*. Cambridge, MA: Harvard University Press.

Hjelde, A. (1992). *Trøndsk talemål i Amerika [The Troender Variety of Norwegian in America]*. Trondheim: Tapir.

Hjelde, A. (1996). "The gender of English nouns used in American English," in *Language Contact Across the North Atlantic*, eds P. S. Ureland and I. Clarkson (Tübingen: Max Niemeyer Verlag), 297–312.

Hockett, C. F. (1958). *A Course in Modern Linguistics*. New York, NY: MacMillan.

Hovdenak, M., Killingbergtrø, L., Lauvhjell, A., Nordlie, S., Rommetveit, M. and Worren, D. (1998). *Nynorskordboka [Nynorsk dictionary]*. Oslo: Det norske samlaget.

Jahr, E. H. (1998). "Sociolinguistics in historical language contact: The Scandinavian languages and Low German during the Hanseatic period," in *Language Change: Advances in Historical Sociolinguistics*, ed E. H. Jahr (Berlin: Mouton de Gruyter), 119–130.

Johannessen, J. B. (2015). "The Corpus of American Norwegian Speech (CANS)," in *Proceedings of the 20th Nordic Conference of Computational Linguistics, NODALIDA 2015*, ed. B. Megyesi (Oslo), 297–300.

Johannessen, J. B., and Laake, S. (2015). "On two myths of the Norwegian language in America: is it old-fashioned? Is it approaching the written Bokmål standard?," in *Germanic Heritage Languages in North America*, ed J. B. Johannessen and J. Salmons (Amsterdam: John Benjamins), 299–322.

Johannessen, J. B., and Larsson, I. (2015). Complexity matters: on gender agreement in Heritage Scandinavian. *Front. Psychol*. 6:1842. doi: 10.3389/fpsyg.2015.01842

Johannessen, J. B., Priestley, J., Hagen, K., Åfarli, T. A., and Vangsnes, Ø. A. (2009). "The Nordic Dialect Corpus - an Advanced Research Tool," in *Proceedings of the 17th Nordic Conference of Computational Linguistics NODALIDA 2009, Vol. 4*," eds K. Jokinen and E. Bick (NEALT Proceedings Series).

Johannessen, J. B., and Salmons, J. (2012). Innledning. [Introduction]. *Norsk Lingvistisk Tidsskrift [Nor. J. Linguist.]* 30, 139–148.

Johannessen, J. B., and Salmons, J. (2015). "The Study of Germanic heritage languages in the Americas," in *Germanic Heritage Languages in North America*, eds J. B. Johannessen and J. Salmons (Amsterdam: John Benjamins), 1–17.

Julien, M. (2005). *Nominal Phrases from a Scandinavian Perspective*. Amsterdam: John Benjamins.

Kürschner, S., and Nübling, D. (2011). The interaction of gender and declension in Germanic languages. *Folia Linguistica* 45, 355–388. doi: 10.1515/flin. 2011.014

Lødrup, H. (2011). Hvor mange genus er det i Oslo-dialekten? [How many genders are there in the Oslo dialect?] *Maal og Minne* 2, 120–136.

Montrul, S. (2002). Incomplete acquisition and attrition of Spanish tense/aspect distinctions in adult bilinguals. *Bilingualism* 5, 39–68. doi: 10.1017/S1366728902000135

Montrul, S. (2008). *Incomplete Acquisition in Bilingualism: Re-examining the Age Factor*. Amsterdam: John Benjamins.

Montrul, S., Foote, R and Perpiñán, S. (2008). Gender agreement in adult second language learners and Spanish Heritage Speakers: the effects of age and context of acquisition. *Lang. Learn*. 58, 503–553. doi: 10.1111/j.1467-9922.2008. 00449.x

Nesse, A. (2002). *Språkkontakt Mellom Norsk og Tysk I Hansatidens Bergen [Language Contact between Norwegian and German in Bergen during the Hanseatic Time]*. Oslo: Novus

Nygård, M., and Åfarli, T. A. (2013). "The structure of gender assignment and American Norwegian," in *Paper Presented at the 4th Annual Workshop on Immigrant Languages in the Americas*, Reykjavík: University of Iceland.

O'Grady, W., Kwak, H-Y., Lee, O-S., and Lee, M. (2011). An emergentist perspective on heritage language acquisition. *Stud. Sec. Lang. Acquis*. 33, 223–245. doi: 10.1017/S0272263110000744

Pascual y Cabo, D., and Rothman, J. (2012). The (Il)logical Problem of Heritage Speaker Bilingualism and incomplete acquisition. *Appl. Linguist*. 33, 450–455. doi: 10.1093/applin/ams037

Polinsky, M. (1997). "American-Russian: language loss meets language acquisition," in *Formal Approaches to Slavic Linguistics*, eds W. Browne, E. Dornisch, N. Kondrashova, and D. Zec (Ann Arbor: Michigan Slavic Publications), 370–407.

Polinsky, M. (2006). Incomplete acquisition: American Russian. *J. Slavic Linguist*. 14, 191–262.

Polinsky, M. (2008). Gender under incomplete acquisition: Heritage speakers' knowledge of noun categorization. *Herit. Lang. J*. 6, 40–71.

Putnam, M., and Sánchez, L. (2013). What's so incomplete about incomplete acquisition? A prolegomenon to modeling heritage language grammars. *Linguist. Approach. Bilingualism* 3, 478–508. doi: 10.1075/lab.3. 4.04put

Rodina, Y., and Westergaard, M. (2013). The acquisition of gender and declension class in a non-transparent system: monolinguals and bilinguals. *Stud. Linguist*. 67, 47–67. doi: 10.1111/stul.12012

Rodina, Y., and Westergaard, M. (2015a). Grammatical gender in Norwegian: Language acquisition and language change. *J. Germ. Linguist*. 27.2, 145–187. doi: 10.1017/S1470542714000245

Rodina, Y., and Westergaard, M. (2015b). Grammatical gender in bilingual Norwegian-Russian acquisition: the role of input and transparency. *Bilingual. Lang. Cogn*. doi: 10.1017/S1366728915000668. [Epub ahead of print].

Rothman, J. (2007). Heritage speaker competence differences, language change, and input type: Inflected infinitives in Heritage Brazilian Portuguese. *Int. J. Bilingualism* 11, 159–389. doi: 10.1177/13670069070110 040201

Rothman, J. (2009). Understanding the nature and outcomes of early bilingualism:Romance languages as heritage languages. *Int. J. Bilingualism* 13, 155–163. doi: 10.1177/1367006909339814

Schmid, M. (2002). First *Language Attrition, Use and Maintenance. The Case of German Jews in Anglophone Countries*. Philadelphia, PA: John Benjamins.

Sorace, A. (2004). Native language attrition and developmental instability at the syntax-discourse interface: data, interpretations and methods. *Bilingual. Lang. Cogn*. 7, 143–145. doi: 10.1017/S13667289040 01543

Trosterud, T. (2001). Genustilordning I norsk er regelstyrt. [Assignment of gender in Norwegian is rule-based]. *Norsk Lingvistisk Tidsskrift [Norwegian J. Ling.]* 19, 29–57.

Trudgill, P. (2013). "Gender maintenance and loss in Totenmålet, English, and other major Germanic varieties," in *In Search of Universal Grammar:*

From Old Norse to Zoque, ed T. Lohndal (Amsterdam: John Benjamins), 77–107.

Tsimpli, I. M., Sorace, A., Heycock, C., and Filiaci, F. (2004). First language attrition and syntactic subjects: a study of Greek and Italian near-native speakers of English. *Int. J. Bilingualism* 8, 257–277. doi: 10.1177/13670069040080030601

Vangsnes, Ø. A. (1999). *The Identification of Functional Architecture*. PhD Dissertation, University of Bergen.

Venås, K. (1993). "On the choice between two written standards in Norway," in *Language Conflict and Language Planning,* ed E. H. Jahr (Berlin; New York, NY: Mouton de Gruyter), 263–278.

Westergaard, M., and Anderssen, M. (2015). "Word Order Variation in Norwegian Possessive Constructions: Bilingual Acquisition and Attrition," in *Germanic heritage languages in North America: Acquisition, attrition and* change, eds J. B. Johannessen and J. Salmons (Amsterdam: John Benjamins), 21–45.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

APPENDIX

TABLE A1 | Production of the indefinite article for each of the three genders by all speakers in CANS (N = 50).

	M			F			N		
Informant	M	F	N	M	F	N	M	F	N
GROUP 1									
harmony_MN_03gm	5								
harmony_MN_05gm	1								
north_battleford_SK_02gk									
spring_grove_MN_09gm	3								
GROUP 2									
billings_MT_01gm	8			2					1
blair_WI_01gm	9			4					
blair_WI_07gm	7			2					
spring_grove_MN_05gm	4			2					
zumbrota_MN_01gk	4			1					1
GROUP 3									
blair_WI_02gm	9			2			2		
blair_WI_04gk	16			4					
coon_valley_WI_01gk	5			2			2		
coon_valley_WI_12gm	3						2		
decorah_IA_02gm	5			2			1		
gary_MN_01gm	14						2		
sunburg_MN_04gk	5			4			1		
vancouver_WA_03uk	4			1			1		
westby_WI_02gm	7						2		
GROUP 4									
albert_lea_MN_01gk	8			4	8		1		1
chicago_IL_01gk	17			9			6		5
coon_valley_WI_02gm	14			1	3		3	1	4
coon_valley_WI_03gm	29			1	8				1
coon_valley_WI_04gm	6				2		2		2
coon_valley_WI_06gm	36	1			5		1	3	7
coon_valley_WI_07gk	5			1	3		1		
decorah_IA_01gm	15			4			1		2
fargo_ND_01gm	15			5	4		3		
flom_MN_01gm	14			1	12		3		4
flom_MN_02gm	19				10			2	3
gary_MN_02gk	16	1		3	8		1	1	10
glasgow_MT_01gm	4			3			2		3
harmony_MN_01gk	16			2	2		1		1
harmony_MN_02gk	12				4			4	1
harmony_MN_04gm	5				1		1	1	1
north_battleford_SK_01gm	1						4		2
portland_ND_01gm	8			4	1		3		1
portland_ND_02gk	6			8			13		1
rushford_MN_01gm	1	1		1	1				

(Continued)

TABLE A1 | Continued

	M			F			N		
Informant	M	F	N	M	F	N	M	F	N
stillwater_MN_01gm	75			13	25		9		20
sunburg_MN_03gm	13			3	3		4		2
sunburg_MN_12gk	1				1		1		
vancouver_WA_01gm	14			8			2		2
wanamingo_MN_04gk	2			1	1		1		1
webster_SD_01gm	24			3			1		1
webster_SD_02gm	6			4	4				2
westby_WI_01gm	41	1			30		2	1	15
westby_WI_03gk	13				4			3	6
westby_WI_05gm	3						3	3	
westby_WI_06gm	9	1		3	1		2		
zumbrota_MN_02gm	4			1	4				

The baseline is the Nynorsk dictionary adjusted for some typical patterns in Eastern Norwegian dialects. Group 1, Gender system unclear; Group 2, Possibly a three-gender system; Group 3, Masculine forms only; Group 4, Mixture of gender forms.

Islands and Non-islands in Native and Heritage Korean

Boyoung Kim[1]* and Grant Goodall[2]

[1] *Department of Asian Studies, University of Texas at Austin, Austin, TX, USA,* [2] *Department of Linguistics, University of California, San Diego, La Jolla, CA, USA*

To a large extent, island phenomena are cross-linguistically invariable, but English and Korean present some striking differences in this domain. English has *wh*-movement and Korean does not, and while both languages show sensitivity to *wh*-islands, only English has island effects for adjunct clauses. Given this complex set of differences, one might expect Korean/English bilinguals, and especially heritage Korean speakers (i.e., early bilinguals whose L2 became their dominant language during childhood) to be different from native speakers, since heritage speakers have had more limited exposure to Korean, may have had incomplete acquisition and/or attrition, and may show significant transfer effects from the L2. Here we examine islands in heritage speakers of Korean in the U.S. Through a series of four formal acceptability experiments comparing these heritage speakers with native speakers residing in Korea, we show that the two groups are remarkably similar. Both show clear evidence for *wh*-islands and an equally clear lack of adjunct island effects. Given the very different linguistic environment that the heritage speakers have had since early childhood, this result lends support to the idea that island phenomena are largely immune to environmental influences and stem from deeper properties of the processor and/or grammar. Similarly, it casts some doubt on recent proposals that islands are learned from the input.

Keywords: island constraints, Korean, heritage speakers, acquisition, scope ambiguity, *wh-in-situ*

INTRODUCTION

A well-known fact about filler-gap dependencies in natural language is that gaps are not allowed in certain structural environments, known as islands. Interrogative clauses (*wh*-clauses or whether-clauses) are one such environment, for instance, as seen in (1).

(1) a. *Who do you wonder [why Mary saw __] ?
    b. *Who do you wonder [whether Mary saw __] ?

One interesting fact about islands is that to a very large extent, they are cross-linguistically invariable. That is, environments where gaps are disallowed in English often have this same characteristic in other languages. Likely related to this is the fact that children's sensitivity to islands does not seem to depend in any obvious way on their being exposed to direct evidence for them. Children clearly hear evidence for filler-gap dependencies and for structures such as *wh*-clauses, for instance, but it is not clear if anything in the environment would suggest to children that gaps should not be allowed within such clauses. For this reason, many have suggested that islands are not learned directly, but instead follow from constraints on processing ability (e.g.,

Kluender, 1998; Hofmeister and Sag, 2010) or grammar (e.g., Chomsky, 1971; Rizzi, 2013). In recent years, however, some have suggested that initial appearances notwithstanding, islands are in fact able to be learned from the environment (e.g., Culicover and Jackendoff, 2005; Pearl and Sprouse, 2013). Under this type of approach, children use statistical mechanisms to track, analyze, and generalize patterns in the input. On the basis of the generalizations attained, they are able to produce and comprehend sentences beyond their experience, while prohibiting patterns not warranted by their experience.

Despite the apparent cross-linguistic uniformity of islands, there is nonetheless some variability. To begin with, many languages do not have overt *wh*-extraction, adopting instead a wh *in-situ* strategy, in which the *wh*-phrase occupies what would otherwise be the gap position. In such languages, there is thus no overt filler-gap dependency. This may be seen in the Korean example in (2), in which the *wh*-phrase nwukwu-ul "who" is located in the embedded clause, and the scope of this phrase is indicated by the question particle –ni in the matrix clause.

(2) Mary -nun [Obama-ka nwukwu-ul manna-ss-ta-ko]
    -Top -Nom   who -Acc   meet-Past-Decl-Comp
    malhae-ss-ni?
    say -Past-Q
         "Who did Mary say that Obama met ____?"

As will be discussed below, the *wh*-phrase here may be also interpreted as an existential pronoun, in which case the question particle –ni would signal a yes/no question, but the important point here is that even under the *wh*-question interpretation, nwukwu-ul remains *in-situ*.

Even in languages of this type, though, it has sometimes been claimed that islands are still obeyed, in the sense that the *in-situ wh*-phrase is degraded when inside an island structure. This has been claimed for interrogative clauses in Japanese and Korean, for example, as seen in (3) for Korean (claimed to be unacceptable under the *wh*-question reading given here).

(3) *Mary –nun [Obama-ka <u>nwukwu</u>-ul manna-ss-nun-ci]
    -Top –Nom   who -Acc   meet-Past-And-Q
    ahl-ass-ni?
    know-Past-Q
         "Who did Mary know whether Obama met ____?"

The picture is not quite this simple, though, since some environments that appear to be islands in *wh*-movement languages nonetheless allow *wh*-phrases in wh *in-situ* languages. Adjunct clauses provide an example of this. (4) shows that they are typically islands in a language like English, while (5) shows that they appear to allow *wh*-phrases in a language like Korean.

(4)   *Who did Mary appear [when Obama met __]?

(5) Mary-nun [Obama-ka **nwukwu**-ul manna-ss-ul-**ttay**]
    -Top -Nom who -Acc   meet-Past-Adn-when
    natana-ss-**ni**?
    appear-Past-Q
         "Who did Mary appear when Obama met?"

What we have seen so far describes the knowledge of monolingual speakers of languages like English and Korean. Bilingual speakers of two languages with these properties would very reasonably be expected to be different. Bilinguals receive less input for each of their languages than do monolinguals for their single language. For heritage speakers in particular (i.e., early bilinguals who grew up with exposure to the heritage language (L1) and the majority language (L2) either simultaneously or sequentially in early childhood, but whose L2 became the primary language at some point during childhood), there are additional factors. Their exposure to the heritage language in childhood may have been limited in important ways, their acquisition of the heritage language may have been incomplete, and they may have undergone significant attrition in the years since childhood (e.g., Anderson, 1999; Montrul, 2002; Sorace, 2004; de Groot, 2005; Polinsky, 2011a). In addition, for all bilinguals, there is the possibility of transfer, that is, that properties of one language will influence the other. For bilinguals in general, and heritage speakers in particular, it is well known that these environmental differences can lead to very significant differences between them and native speakers (e.g., Polinsky, 2011b and references cited there).

Bilinguals thus present an especially interesting case with regard to island phenomena. On the one hand, the various environmental differences just described, along with the possibility of transfer, could very reasonably be expected to lead to differences between bilinguals and monolinguals, especially with languages like English and Korean, where the island properties are so different. On the other hand, many have suggested that islands arise not because of learning *per se*, but because of resource limitations on the processor or computational limitations on the grammar, and if such is the case, we would expect few differences between monolinguals and bilinguals, assuming that their processing and/or grammatical resources are similar.

In this paper, we examine islands in heritage speakers of Korean in the United States, i.e., early Korean/English bilinguals for whom English has become the dominant language. If islands are susceptible to environmental influences, then there are many reasons to expect this type of bilingual to behave differently with regard to islands than monolingual speakers, as we have seen. If, on the other hand, islands are primarily the result of specific properties of the processor and/or grammar, then we would expect these bilinguals to display island behavior that is basically the same as monolinguals.

We will focus in particular on the heritage speakers' sensitivity to islands in Korean. As we have seen, Korean is a *wh in-situ* language and has been claimed to show island effects in *wh*-clauses, but not in adjunct clauses. English, on the other hand, has *wh*-movement and shows island effects in both *wh*-clauses

and adjunct clauses. We perform the same set of formal acceptability experiments on native Korean speakers residing in Korea and on heritage speakers of Korean living in the U.S. As we will see, our experiments show that sensitivity to islands is very similar in the two groups, lending support to the idea that island phenomena are largely immune to environmental influences and stem from deeper properties of the processor and/or grammar.

The paper is structured as follows: Section Island Effects in Korean gives further details about the nature of island constraints in Korean, Section Experiments presents a series of four experiments probing this phenomenon in both native and heritage speakers, and Section Conclusion presents the overall conclusions.

ISLAND EFFECTS IN KOREAN

Given the lack of overt *wh*-movement in the language, the question of whether a given island effect obtains in Korean reduces to the question of whether it is possible for an *in-situ* *wh*-phrase within a putative island domain to take scope outside of that domain. For adjunct clauses, it is usually thought that such wide-scope readings are in fact possible, as in (5) above, and for this reason, adjunct clauses are often believed not to have island status in Korean. For *wh*-clauses, however, the facts are not as clear. Many have claimed that a wide-scope reading for a *wh*-phrase in such clauses, as in (3) above, is not possible (e.g., Lee, 1982; Kim, 1989; Nishigauchi, 1990; Han, 1992; Watanabe, 1992 for Japanese; Hong, 2004 for Korean), suggesting that these clauses are islands, but others have claimed that it is possible (e.g., Suh, 1987; Ishihara, 2002; Choi, 2006; Hwang, 2007 for Korean; Sprouse et al., 2011 for Japanese), suggesting that these clauses are not islands.

One of the reasons for this lack of clarity surrounding the status of *wh*-islands in the literature is the fact that simple acceptability judgments of the string are not sufficient to decide the matter. First, the issue is how the scope of the *wh*-phrase is interpreted, not whether the sentence is acceptable or not. (3) is uncontroversially acceptable, for instance, if one gives a narrow-scope reading to the *wh*-word, as in (6).

(6)   Did Mary know who Obama met ___?

Second, in addition to their interpretation as interrogatives, bare *wh*-words in Korean may also be interpreted as existential pronouns. Thus, in addition to the two readings already given for (3), an interpretation as in (7) is also possible.

(7)   Did Mary know whether Obama met someone?

Combined, these two facts mean that for any given question with a *wh*-word, if one reading is not available, another typically is, with the result that all such questions give the appearance of being acceptable. This has made exploring the possibility of island effects in Korean very difficult and has no doubt contributed to the lack of consensus in the literature regarding *wh*-islands in Korean. In this study, we are able to circumvent this problem by presenting participants with question-answer pairs and soliciting acceptability judgments on the answer, rather than the question.

Given that the answer will be appropriate for one reading of the question but not others, we are thus able to obtain, albeit indirectly, an acceptability rating for a particular reading. Using this technique, we are able to accomplish two goals. First, we are able to establish clearly the extent to which island effects exist in Korean, despite the lack of clarity in the literature. Second, we are able to make precise comparisons in this regard between native speaker controls and heritage speakers.

EXPERIMENTS

The following four experiments use the technique just described to explore the possibility of island effects in Korean. We test both *wh*-clauses and adjunct clauses, using both native Korean speaker controls and heritage speakers of Korean (Korean/English bilinguals).

Experiment 1: Canonical *Wh*-Islands in Korean

Participants

Twenty-eight English-dominant heritage speakers of Korean, all students at UCSD, participated for course credit. Thirty three percent of the heritage participants were US-born and 67% were Korean-born and moved to the U.S. from Korea before age 7 (*M:* 3 years old, *SD*: 2.7). Their mean age at the time of testing was 20 (range: 18–25, *SD*: 1.8). Fifty seven percent of the heritage speakers reported that Korean was their mother tongue, 33% reported English, and the remaining 10% reported both languages. 86% of the parents spoke only Korean with them, and 14% spoke both languages. All were literate in Korean. As a control group, 48 native speakers of Korean who were residing in Korea at the time of testing participated online (*M*: 28 years old, range: 20–34, *SD*: 3.7).

After the experiment, participants took a Korean proficiency test. The proficiency test consisted of a cloze test, and multiple choice questions on synonym-antonym. The proficiency test results indicated that heritage speakers (*M*: 78%, range: 50–100%, *SD*: 16.7) were significantly less proficient than native speakers (*M*: 96%, range: 88–100%, *SD*: 3.1) [$F_{(1, 74)} = 59.1, p < 0.0001$].

Stimuli

Since island effects in Korean can be tested only by examining speakers' interpretation of sentences (i.e., *wh*-scope), we will measure the felicity of Question-Answer pairs. Variants of this method have been used in several studies testing scope ambiguity of *wh-in-situ* (e.g., Pesetsky, 1987; Umeda, 2008; Kitagawa and Hirose, 2012). The specifics of the experimental design are as follows.

We present participants a set of a context, a question (containing an island configuration), and an answer. Then, instead of asking for the acceptability of the question, we ask them to rate the acceptability of the answer as a very first response to the *wh*-question. The answers consist of two types: either "*wh*-answers" or "yes/no answers," "*Wh*-answers" are appropriate for a direct *wh*-question interpretation of the preceding question, while "yes/no answers" are appropriate for a yes/no question interpretation. The answers would thus

encourage one reading or the other. The acceptability of *wh*-answers would reflect the possibility of the island-violating interpretation when a *wh*-word is interpreted as a *wh*-question word with scope outside the embedded clause. On the other hand, when the *wh*-word is interpreted as an indefinite pronoun, or as a true *wh*-word with scope over only the embedded clause (yielding an indirect question), a yes/no question results.

There were thus three factors (Location of *wh*-word, Structure of embedded clause, Answer type), with a total of eight conditions. Stimuli consisted of question-answer pairs, preceded by a context. All question sentences were biclausal. As we will see below, there is optionality in the position of embedded clauses in Korean, but in this experiment, all embedded clauses immediately precede the matrix verb. They differed as to the *Location* of the *wh*-word (matrix vs. embedded clause) and the *Structure* of the embedded clause [declarative (non-island) vs. interrogative (island)]. There were also two different types of answers, either "*wh*-answers" or "yes/no answers." Sample stimuli are provided in (8)–(15). In (8)–(9), the *wh*-word is in the matrix clause and the embedded clause is declarative, while in (10)–(11), the embedded clause is interrogative. In (12)–(13), the *wh*-word is in an embedded clause that is declarative, while in (14)–(15), the embedded clause is interrogative.

(8) Q: **Nwukwu**-ka [Obama-ka Mary-ul   manna-ss-ta-**ko**]
       who -Nom -Nom -Acc   meet-Past-Decl-that
       tul-ess-**ni**?
       hear-Past-Q
       "Who heard that Obama met Mary?" or
       "Did somebody hear that Obama met Mary?"
    A: *WH*-ANSWER: Hillary-ka "Hillary"

(9) Q: Same as (8).
    A: YES-NO ANSWER: Ney, tul-ess-eyo "Yes, heard"

(10) Q: **Nwukwu**-ka [Obama-ka Mary-ul   manna-ss-nun-**ci**]
        who -Nom -Nom -Acc   meet-Past-Adn-Q
        tul-ess-**ni**?
        hear-Past-Q
        "Who heard whether Obama met Mary?" or
        "Did somebody hear whether Obama met Mary?"
     A: *WH*-ANSWER: Hillary-ka "Hillary"

(11) Q: Same as (10).
     A: YES-NO ANSWER: Ney, tul-ess-eyo "Yes, heard"

(12) Q: Mary-nun [Obama-ka **nwukwu**-ul   manna-ss-ta-**ko**]
        -Top -Nom who -Acc   meet-Past-Decl-that
        tul-ess-**ni**?
        hear-Past-Q
        "Who did Mary hear that Obama met?" or
        "Did Mary hear that Obama met somebody?"
     A: *WH*-ANSWER: Hillary-lul "Hillary"

(13) Q: Same as (12).
     A: YES-NO ANSWER: Ney, tul-ess-eyo "Yes, heard"

(14) Q: Mary-nun [Obama-ka **nwukwu**-ul   manna-ss-nun-**ci**]
        -Top -Nom who -Acc   meet-Past-Adn-Q
        tul-ess-**ni**?
        hear-Past-Q
        "Who did Mary hear whether Obama met?" or
        "Did Mary hear who Obama met?"
     A: *WH*-ANSWER: Hillary-lul "Hillary"

(15) Q: Same as (14).
     A: YES-NO ANSWER: Ney, tul-ess-eyo "Yes, heard"

All question-answer pairs were preceded by a context consisting of a situation (e.g., "at the White House") and a list of people involved in the situation (e.g., "Mary, Obama, Hillary"). These contexts were designed to make the *wh*-answer pragmatically plausible, even when this interpretation of the question would violate an island. All experimental stimuli were in Korean, but the English translation was also provided for the context part for the heritage speakers.

Forty sets of experimental sentences were distributed using a Latin Square design among eight lists consisting of five tokens of each of the eight conditions. Each list included 63 fillers, for an experimental/filler ratio of 1:1.5. All fillers were questions, some with and some without a *wh*-pronoun, representing a wide range of acceptability. All lists were randomized.

In 30 of the 40 sets, the matrix verb was matched across all conditions in the set. In the remaining 10 sets, however, one verb is used with declarative complements and another verb with interrogative complements (e.g., *sayngkakhata* "think" with declaratives and *kungkumhata* "wonder" with interrogatives. This was due to the limited number of verbs (e.g., *tutta* "hear") that can take both declarative and interrogative complements. The *wh*-word *nwukwu* "who" was used in all stimuli.

All stimuli were presented in written form and thus without an explicit indication of prosody. It appears that prosody is able to ameliorate some possible island effects in Japanese and Korean (e.g., Kitagawa, 2005), but it may not be able to eliminate them entirely (e.g., Hwang, 2007). In any event, since our goal here is to compare native speakers and heritage speakers, what matters is that the experimental stimuli be identical, and that condition is met.

Method

The experiments were conducted in the Experimental Syntax Lab at UCSD for heritage speakers, and online for native speakers. The experiments in this study were approved by the Institutional Review Board of the University of California, San Diego (#110080). All subjects involved gave their informed written consent. Subjects were instructed to rate the acceptability of the answer as a first response to the question, using a 7-point scale (with 1 "very bad" and 7 "very good"). A sample item is given in **Figure 1**.

Analysis

Acceptability scores from each participant were z-score transformed prior to analysis, and a series of repeated-measures ANOVAs were conducted on the z-score results. Each group's
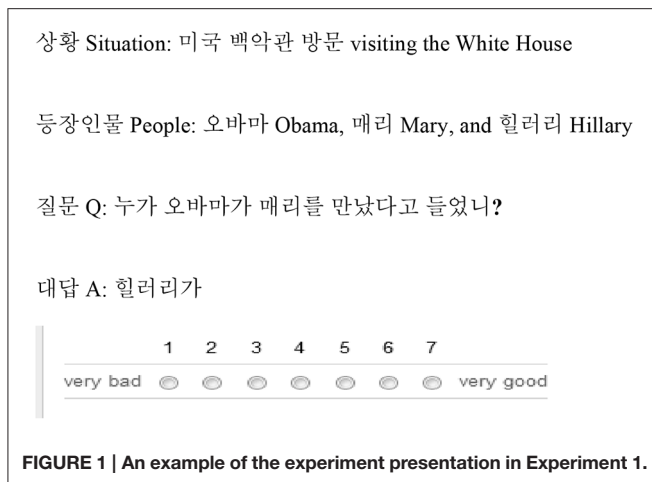
FIGURE 1 | An example of the experiment presentation in Experiment 1.

FIGURE 2 | Results of Experiment 1.

data were separated by answer type, and separate repeated measures ANOVAs were run for each answer type in each group, with Location of *wh*-word (matrix vs. embedded) and Structure of embedded clause (non-island "declarative" vs. island "interrogative") as within-subjects variables, and "subject" (F1) and "item" (F2) as random factors.

An interaction between Location and Structure, where the embedded *wh*-word in an interrogative clause is of lower acceptability than the other three conditions, will be suggestive of an island effect. In order to compare the effect size between groups of any such interaction, differences-in-differences scores (DD) are calculated as follows for each participant using the z-scores for the *wh*-answer type: DD = D1 (Non-Island/Embedded—Island/Embedded)—D2 (Non-Island/Matrix—Island/Matrix). A positive DD score signals super-additivity: the result is more than the sum of the two individual experimental factors. A larger DD score represents a larger island effect, while a negative DD score represents a sub-additive (non-island) interaction.

Results

The results are plotted in **Figure 2** (error bars in all figures represent SE). The first two graphs are natives' results and the following two graphs are heritage speakers'. In both groups, the left graph represents the acceptability of *wh*-answers and the right graph shows that of yes/no answers.

First, with *wh*-answers, in the results of both groups, when a *wh*-word is located in the matrix clause, the two types of structures were rated similarly, but with an embedded *wh*-word, the declarative condition was preferred over the interrogative condition, indicating dispreference for the matrix *wh*-scope of the embedded *wh*-word, that is the *wh*-island effect. This was also shown by significant main effects for Location [native: $F1_{(1, 47)} = 10.17$, $p = 0.003$, $F2_{(1, 39)} = 15.22$, $p < 0.0001$; heritage: $F1_{(1, 27)} = 27.66$, $p < 0.0001$, $F2_{(1, 39)} = 44.14$, $p < 0.0001$], and Structure [native: $F1_{(1, 47)} = 29.83$, $p < 0.0001$, $F2_{(1, 39)} = 28.12$, $p < 0.0001$; heritage: $F1_{(1, 27)} = 48.86$, $p < 0.0001$, $F2_{(1, 39)} = 32.57$, $p < 0.0001$]. The interaction between these two factors was significant for natives [$F1_{(1, 47)} = 16.17$, $p < 0.0001$,
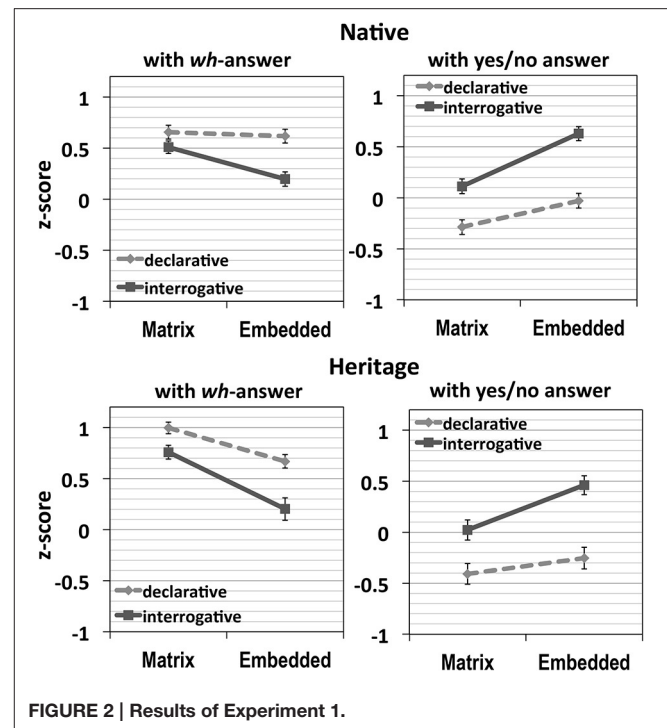
$F2_{(1, 39)} = 7.13$, $p = 0.011$], and marginal for heritage speakers [$F1_{(1, 27)} = 3.47$, $p = 0.07$, $F2_{(1, 39)} = 3.41$, $p = 0.07$].

The differences-in-differences (DD) scores in both groups were positive [Native: 0.28 (*SD*: 0.48), Heritage: 0.23 (*SD*: 0.65)], indicating a super-additive *wh*-island effect in both groups. A one-way ANOVA with DD-score as a dependent factor, and Group as a fixed factor yielded no significant difference between the two groups ($p = 0.71$).

With yes/no answers, the pattern was reversed, with higher acceptability with embedded *wh*-words, than with matrix *wh*-words. Crucially, the condition with an embedded *wh*-word inside a *wh*-clause was preferred to be answered with yes/no answers, more than in any other conditions, indicating a wh-island effect. Both groups displayed main effects for Location [native: $F1_{(1, 47)} = 33.64$, $p < 0.0001$, $F2_{(1, 39)} = 39.21$, $p < 0.0001$; heritage: $F1_{(1, 27)} = 18.08$, $p < 0.0001$, $F2_{(1, 39)} = 19.74$, $p < 0.0001$], and Structure [native: $F1_{(1, 47)} = 76.08$, $p < 0.0001$, $F2_{(1, 39)} = 61.83$, $p < 0.0001$; heritage: $F1_{(1, 27)} = 54.66$, $p < 0.0001$, $F2_{(1, 39)} = 71.96$, $p < 0.0001$]. In addition, for natives, the interaction of Location and Structure was significant in the subjects analysis and close to significant in the items analysis [$F1_{(1, 47)} = 5.04$, $p = 0.03$, $F2_{(1, 39)} = 3.89$, $p = 0.056$], while for heritage speakers, the interaction approached significance in both types of analysis [$F1_{(1, 27)} = 3.89$, $p = 0.059$, $F2_{(1, 39)} = 3.15$, $p = 0.08$].

In sum, these results suggest a very clear *wh*-island effect in Korean for the natives. That is, when the *wh*-word is located within an embedded interrogative clause, the *wh*-answer is strongly dispreferred and a yes/no answer is strongly preferred. Since the *wh*-answer is only compatible with matrix scope for the *wh*-word and the yes/no answer is only compatible with

embedded scope, these results suggest that the *wh*-word is not able to scope out of the embedded interrogative clause. For heritage speakers, the situation is less clear. They exhibit a numerically similar pattern suggestive of a *wh*-island effect, but this effect does not reach significance. We return to this issue in Experiment 3, where we test for the existence of a *wh*-island effect in the two populations by means of stimuli where the interrogative clause is in sentence-initial position, as is also possible in Korean.

Experiment 2: Acceptability of Canonical Adjunct-Islands in Korean
Participants, Method, and Analysis

The participants, method, and analysis of the results were the same as in Experiment 1.

Stimuli

The basic design of the experiment is the same as in Experiment 1, consisting of a total of 8 conditions, reflecting three factors: Location of *wh*-word (matrix vs. embedded) × Structure of embedded clause [complement (non-island) vs. adjunct (island)] × Answer type (*wh*-answer vs. yes/no-answer). What distinguishes this experiment from the previous one is that here we are contrasting embedded complement clauses with embedded adjunct clauses. As in Experiment 1, all embedded clauses immediately precede the verb here, although other positions are also possible (see Experiments 3 and 4 below).

All 8 conditions in this experiment were lexically matched except for the matrix verb, which had to differ between complement clauses and adjunct clauses for selectional reasons (e.g., *tutta* "hear" in complement conditions vs. *natanata* "appear" in adjunct conditions).

As in Experiment 1, 40 sets of experimental sentences were distributed using a Latin Square design among eight lists consisting of five tokens of each of the eight conditions. Each list included 63 fillers, for an experimental/filler ratio of 1:1.5. All lists were randomized. The *wh*-word *nwukwu* "who" was used in all stimuli. Sample stimuli are provided in (16)–(23).

(16) Q: **Nwukwu**-ka [Obama-ka Mary-ul   manna-ss-ta-**ko**]
        who -Nom   -Nom -Acc   meet-Past-Decl-that
        tul-ess-**ni**?
        hear-Past-Q
        "Who heard that Obama met Mary?" or
        "Did somebody hear that Obama met Mary?"
    A: *WH*-ANSWER: Hillary-ka "Hillary"

(17) Q: Same as (16).
    A: YES-NO ANSWER: Ney, tul-ess-eyo "Yes, heard"

(18) Q: **Nwukwu**-ka [Obama-ka Mary-ul   manna-ss-ul-**ttay**]
        who -Nom   -Nom -Acc   meet -Past-Adn-when
        natana-ss-**ni**?
        appear-Past-Q
        "Who appeared when Obama met Mary?" or
        "Did somebody appear when Obama met Mary?"
    A: *WH*-ANSWER: Hillary-ka "Hillary"

(19) Q: Same as (18).
    A: YES-NO ANSWER: Ney, natana-ss-eyo "Yes, appeared"

(20) Q: Mary-nun [Obama-ka **nwukwu**-ul   manna-ss-ta-**ko**]
        -Top -Nom who -Acc   meet-Past-Decl-that
        tul-ess-**ni**?
        hear-Past-Q
        "Who did Mary hear that Obama met?" or
        "Did Mary hear that Obama met somebody?"
    A: *WH*-ANSWER: Hillary-lul "Hillary"

(21) Q: Same as (20).
    A: YES-NO ANSWER: Ney, tul-ess-eyo "Yes, heard"

(22) Q: Mary-nun [Obama-ka **nwukwu**-ul manna-ss-ul-**ttay**]
        -Top -Nom who -Acc   meet-Past-Adn-when
        natana-ss-**ni**?
        appear-Past-Q
        "Who did Mary appear when Obama met?" or
        "Did Mary appear when Obama met somebody?"
    A: *WH*-ANSWER: Hillary-lul "Hillary"

(23) Q: Same as (22).
    A: YES-NO ANSWER: Ney, natana-ss-eyo "Yes, appeared"

Results

In **Figure 3**, the first two graphs represent natives', and the following two graphs are heritage speakers' results. In each set of graphs, the first graph shows the results with the *wh*-answer, and the second graph displays the results with the yes/no answer.

The acceptability of the adjunct clause conditions did not change much depending on the location of the *wh*-word with both types of answers in both groups, indicating the absence of
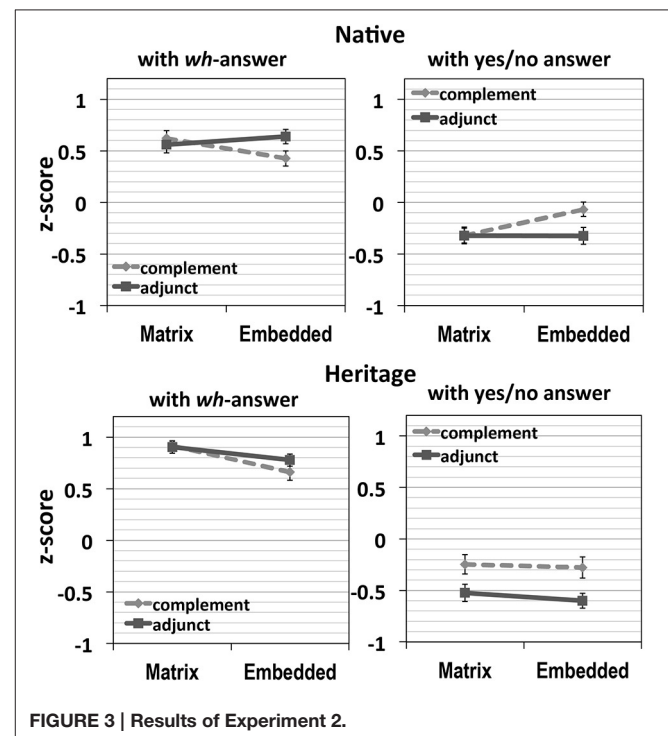

FIGURE 3 | Results of Experiment 2.

adjunct island effects. First, for the heritage speakers, a *wh*-word within an adjunct clause does not result in significantly decreased acceptability with *wh*-answers or increased acceptability with yes/no answers, as may be seen in the lack of an interaction between Structure and Location [with *wh*-answer: $F1_{(1, 27)} = 1.49$, $p = 0.23$, $F2_{(1, 39)} = 1.61$, $p = 0.21$; with yes/no answer: $F1_{(1, 27)} = 0.14$, $p = 0.71$, $F2_{(1, 39)} = 0.19$, $p = 0.66$].

The results are similar for the native speakers in that there is no evidence of any adjunct island effect. However, the native group showed a main effect of Structure on the yes/no answers [$F1_{(1, 47)} = 8.52$, $p = 0.005$, $F2_{(1, 39)} = 8.25$, $p = 0.007$], as well as a mostly significant interaction of Structure and Location with both types of answers [with *wh*-answer: $F1_{(1, 47)} = 12.05$, $p = 0.001$, $F2_{(1, 39)} = 4.54$, $p = 0.039$; with yes/no answer: $F1_{(1, 47)} = 6.31$, $p = 0.016$, $F2_{(1, 39)} = 3.04$, $p = 0.089$]. Nevertheless, the direction of the interaction was the opposite of what one would expect for a classic island effect: the condition in which the *wh*-word is located within an adjunct clause was rated the highest out of the four conditions with *wh*-answers, and the lowest with yes/no answers. There is thus no sign of an adjunct island effect for this group.

The differences-in-differences (DD) scores with *wh*-answer were also negative in both groups [native controls: −0.28 (*SD*: 0.56), heritage speakers: −0.13 (*SD*: 0.57)], with no significant difference between the groups. This confirms again no super-additive adjunct island effects in Korean for both groups.

In sum, the reverse interaction of Location and Structure in the native group and the absence of interaction in the heritage group thus very strongly suggest that there are no adjunct island effects in Korean for either group of speakers.

Interim Summary

In Experiments 1 and 2 with canonically ordered embedded interrogative and adjunct clauses, we found *wh*-island effects, but no adjunct island effects in Korean. The *wh*-island violating condition in Experiment 1 was the least acceptable compared to other conditions, while the adjunct island violating condition was rated similarly with its counterparts. The results of the native and heritage groups were similar, thus suggesting that the development of (non-)island effects is largely independent of the learning environment.

In Experiments 3 and 4, we will attempt to replicate these results with different groups of participants and different types of stimuli. The embedded clauses in these experiments will be scrambled to a sentence-initial position. Since this is a natural position for embedded clauses in Korean, and the preferred position for adjunct clauses, it is possible that this will allow for a fairer test for the presence of island effects.

Experiment 3: Acceptability of Scrambled *Wh*-Islands in Korean

Participants

Nineteen English-dominant heritage speakers of Korean, all students at UCSD, participated for course credit. 27% of the heritage participants were US-born and 73% were Korean-born and moved to the U.S. from Korea before age 7 (*M*: 3 years old, *SD*: 2.7). Their mean age at the time of testing was 20 (range: 19–23, *SD*: 1.2). 53% of the heritage speakers reported that Korean was their mother tongue, 21% reported English, and the remaining 26% reported both languages. 85% of the parents spoke only Korean with them, and 15% spoke both languages. 48 native speakers of Korean residing in Korea served as a control group (*M*: 26 years old, range: 20–37, *SD*: 4.8).

After the experiment, participants took the Korean proficiency test, the same one used in Experiments 1 and 2. The proficiency test results implied that heritage speakers (*M*: 78%, range: 51–94%, *SD*: 13.6) were significantly less proficient than native speakers (*M*: 96%, range: 88–100%, *SD*: 3.6) [$F_{(1, 65)} = 76.2$, $p < 0.0001$].

Stimuli, Method, and Analysis

The stimuli differed from those in Experiment 1 only by the location of the embedded clauses: the embedded clauses in this experiment were sentence-initial, whereas those in Experiment 1 were in their canonical (center-embedded) position. There were 8 experimental conditions reflecting 3 factors, just as in Experiment 1: Location of *wh*-word (matrix clause vs. embedded clause) × Structure of embedded clause (declarative vs. interrogative) × Answer type (*wh*-answer vs. yes/no-answer). Sample stimuli are provided in (24)–(31). The methods and analysis of the results were the same as in Experiment 1.

(24) Q: [Obama-ka Mary-ul   manna-ss-ta-**ko**]   **nwukwu**-ka
     -Nom   -Acc   meet-Past-Decl-that   who -Nom
     tul-ess-**ni**?
     hear-Past-Q
     "Who heard that Obama met Mary?" or
     "Did somebody hear that Obama met Mary?"
     A: *WH*-ANSWER: Hillary-ka "Hillary"

(25) Q: Same as (24).
     A: YES-NO ANSWER: Ney, tul-ess-eyo "Yes, heard"

(26) Q: [Obama-ka Mary-ul   manna-ss-nun-**ci**]   **nwukwu**-ka
     -Nom   -Acc   meet-Past-Adn-Q   who -Nom
     tul-ess-**ni**?
     hear-Past-Q
     "Who heard whether Obama met Mary?" or
     "Did somebody hear whether Obama met Mary?"
     A: *WH*-ANSWER: Hillary-ka "Hillary"

(27) Q: Same as in (26).
     A: YES-NO ANSWER: Ney, tul-ess-eyo "Yes, heard."

(28) Q: [Obama-ka **nwukwu**-ul   manna-ss-ta-**ko**]   Mary-ka
     -Nom who -Acc   meet-Past-Decl-that   -Nom
     tul-ess-**ni**?
     hear-Past-Q
     "Who did Mary hear that Obama met?" or
     "Did Mary hear that Obama met somebody?"
     A: *WH*-ANSWER: Hillary-lul "Hillary"

(29) Q: Same as (28).
    A: YES-NO ANSWER: Ney, tul-ess-eyo "Yes, heard"

(30) Q: [Obama-ka **nwukwu**-ul   manna-ss-nun-**ci**]   Mary-ka
      -Nom who -Acc   meet-Past-Adn-Q   -Nom
      tul-ess-**ni**?
      hear-Past-Q
      "Who did Mary hear whether Obama met?" or
      "Did Mary hear who Obama met?"
    A: *WH*-ANSWER: Hillary-lul "Hillary"

(31) Q: Same as in (30).
    A: YES-NO ANSWER: Ney, tul-ess-eyo "Yes, heard."

Results

Similar to the results in Experiment 1 on the *wh*-island effect with a canonically ordered interrogative clause, results in Experiment 3, presented in **Figure 4**, showed the *wh*-island effect with a sentence-initial interrogative clause in both native and heritage groups, but the effect was more robust in Experiment 3. In the results with *wh*-answer, there was no effect of the complement clause type when the *wh*-word is located in the matrix clause, in that all questions with a matrix *wh*-word were rated similarly regardless of the types of embedded clauses.

On the other hand, with an embedded *wh*-word, the island condition was significantly less preferred than the declarative condition. Also, the questions with an interrogative clause showed a distinctive acceptability depending on the location of the *wh*-word, that is the island violating condition was much less acceptable than its counterpart. This all suggests the *wh*-island effect in Korean, which is also supported by the statistical results as in the following.

FIGURE 4 | Results of Experiment 3.

First, natives exhibited main effects of Location [with *wh*-answers $F1_{(1, 47)} = 183.01$, $p < 0.0001$, $F2_{(1, 39)} = 260.41$, $p < 0.0001$]; with yes/no answers [$F1_{(1, 47)} = 85.11$, $p < 0.0001$, $F2_{(1, 39)} = 167.63$, $p < 0.0001$], and Structure [with *wh*-answers $F1_{(1, 47)} = 48.57$, $p < 0.0001$, $F2_{(1, 39)} = 63.24$, $p < 0.0001$]; with yes/no answers [$F1_{(1, 47)} = 28.67$, $p < 0.0001$, $F2_{(1, 39)} = 29.80$, $p < 0.0001$], as well as a significant interaction of Location and Structure [with *wh*-answers, $F1_{(1, 47)} = 42.46$, $p < 0.0001$, $F2_{(1, 39)} = 42.15$, $p < 0.0001$; with yes/no answers, $F1_{(1, 47)} = 6.12$, $p = 0.017$, $F2_{(1, 39)} = 5.86$, $p = 0.02$].

Heritage speakers displayed very similar results, showing main effects of Location [with *wh*-answers $F1_{(1, 18)} = 59.53$, $p < 0.0001$, $F2_{(1, 39)} = 68.70$, $p < 0.0001$]; with yes/no answers [$F1_{(1, 18)} = 87.09$, $p < 0.0001$, $F2_{(1, 39)} = 67.79$, $p < 0.0001$], and Structure [with *wh*-answers $F1_{(1, 18)} = 48.64$, $p < 0.0001$, $F2_{(1, 39)} = 47.29$, $p < 0.0001$]; with yes/no answers [$F1_{(1, 18)} = 101.65$, $p < 0.0001$, $F2_{(1, 39)} = 34.28$, $p < 0.0001$], as well as a significant interaction of Location and Structure [with *wh*-answers, $F1_{(1, 18)} = 26.33$, $p < 0.0001$, $F2_{(1, 39)} = 42.15$, $p < 0.0001$; with yes/no answers, $F1_{(1, 18)} = 17.30$, $p = 0.002$, $F2_{(1, 39)} = 12.45$, $p = 0.001$].

The two groups' island effect size with *wh*-answers, indicated by the differences-in-differences (DD) scores, were very similar to each other [native: 0.71 (SD: 0.75), heritage: 0.72 (SD: 0.61)].

The significant interaction between Location and Structure suggests a strong *wh*-island effect in Korean for both groups. When the *wh*-word is within an embedded interrogative clause, acceptability drops for the *wh*-answer and rises for the yes/no answer, as we would expect if the *wh*-word is unable to take scope out of that clause.

Experiment 4: Acceptability of Scrambled Adjunct-Islands in Korean

Participants

The participants in this experiment were the same as in Experiment 3.

Stimuli, Method, and Analysis

The stimuli in this experiment were the same as those in Experiment 2, but with sentence-initial embedded clauses. There was a total of 3 factors with 8 conditions: Location of *wh*-word (matrix clause vs. embedded clause) × Structure of embedded clause (complement vs. adjunct) × Answer type (*wh*-answer vs. yes/no-answer). Sample stimuli are presented in (32)–(39). The method and analysis were identical to Experiment 2.

(32) Q: [Obama-ka Mary-ul   manna-ss-ta-**ko**]   **nwukwu**-ka
      -Nom -Acc   meet-Past-Decl-that   who -Nom
      tul-ess-**ni**?
      hear-Past-Q
      "Who heard that Obama met Mary?" or
      "Did somebody hear that Obama met Mary?"
    A: *WH*-ANSWER: Hillary-ka "Hillary"

(33) Q: Same as (32).
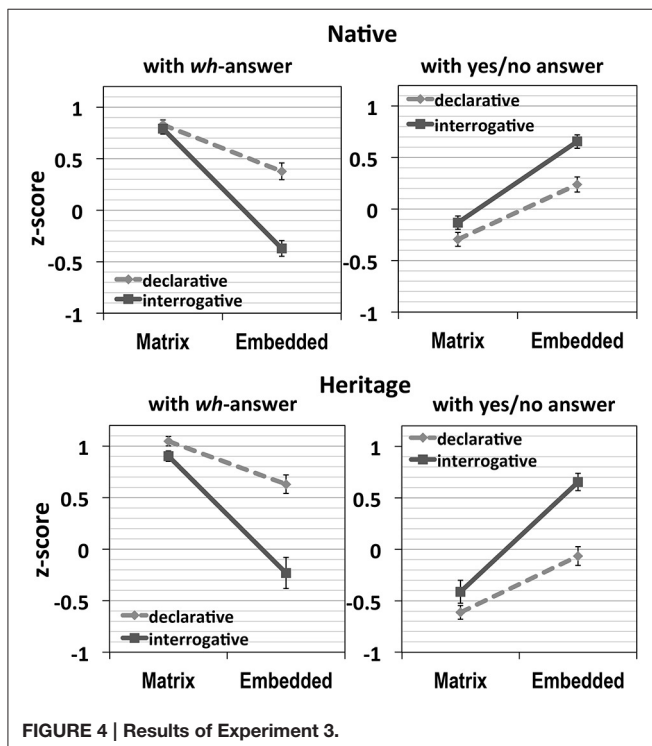   A: YES-NO ANSWER: Ney, tul-ess-eyo "Yes, heard"

(34) Q:[Obama-ka Mary-ul   manna-ss-ul-**ttay**]   **nwukwu**-ka
    -Nom -Acc   meet-Past-Adn-when   who -Nom
    natana-ss-**ni**?
    appear-Past-Q
    "Who appeared when Obama met Mary?" or
    "Did somebody appear when Obama met Mary?"
    A: *WH*-ANSWER: Hillary-ka "Hillary"

(35) Q: Same as in (34).
    A: YES-NO ANSWER: Ney, natana-ss-eyo "Yes, appeared"

(36) Q:[Obama-ka **nwukwu**-ul   manna-ss-ta-**ko**]   Mary-ka
    -Nom who -Acc   meet-Past-Decl-that   -Nom
    tul-ess-**ni**?
    hear-Past-Q
    "Who did Mary hear that Obama met?" or
    "Did Mary hear that Obama met somebody?"
    A: *WH*-ANSWER: Hillary-lul "Hillary"

(37) Q: Same as in (36).
    A: YES-NO ANSWER: Ney, tul-ess-eyo "Yes, heard"

(38) Q:[Obama-ka **nwukwu**-ul   manna-ss-ul-**ttay**]   Mary-ka
    Nom who -Acc   meet-Past-Adn-when   -Nom
    natana-ss-**ni**?
    appear-Past-Q
    "Who did Mary appear when Obama met?" or
    "Did Mary appear when Obama met somebody?"
    A: *WH*-ANSWER: Hillary-lul "Hillary"

(39) Q: Same as in (38).
    A: YES-NO ANSWER: Ney, natana-ss-eyo "Yes, appeared"

Results

As plotted in **Figure 5**, no adjunct island effect was found in either group. Both complement and adjunct clauses received similar acceptability. First, native speakers showed a significant main effect of Location with both *wh*-answers [$F1_{(1, 47)} = 35.02$, $p < 0.0001$, $F2_{(1, 39)} = 40.09$, $p < 0.0001$] and yes/no answers [$F1_{(1, 47)} = 39.79$, $p < 0.0001$, $F2_{(1, 39)} = 47.91$, $p < 0.0001$]. Heritage speakers also revealed a main effect of Location, but the effect was significant only with *wh*-answers [$F1_{(1, 18)} = 10.28$, $p = 0.005$, $F2_{(1, 39)} = 27.99$, $p < 0.0001$] and marginal with yes/no answers [$F1_{(1, 18)} = 3.26$, $p = 0.088$, $F2_{(1, 39)} = 3.99$, $p = 0.053$]. Crucially, neither a main effect of Structure nor an interaction between Location and Structure was significant with either answer type for either group. The differences-in-differences (DD) scores with *wh*-answers were very close to zero in both groups [native: -0.06 (*SD*: 0.78), heritage: -0.09 (*SD*: 0.57)].

The results here provide further support for the conclusion reached in Experiment 2 that there are no adjunct island effects in Korean for either group. The lack of an interaction between Location and Structure suggests that there is no restriction on *wh*-words in adjunct clauses taking wide scope, i.e., that there is no adjunct island.

Summary of the Results in Experiments 1–4

Statistical results of *wh*-answers in Experiments 1–4 are summarized in **Table 1**. As mentioned in Section Stimuli, the results of *wh*-answers reflect the acceptability of the direct
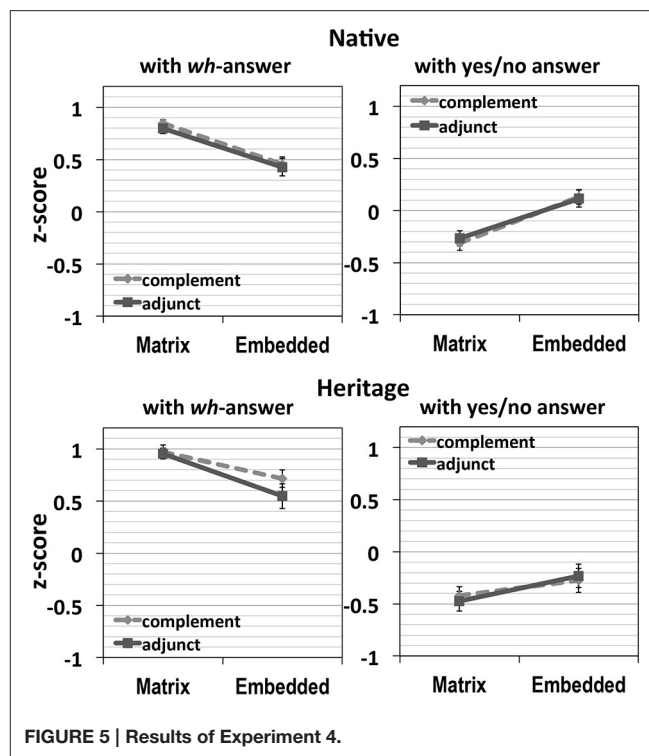

FIGURE 5 | Results of Experiment 4.

wh-question reading where all the *wh*-words are interpreted as *wh*-question words. On the other hand, the results of yes/no answers, specifically with *that*-clauses, indicate the preferred reading of a *wh*-word, either as a question word or as an existential pronoun (i.e., someone) with a *that*-complement clause, while with an interrogative clause, yes/no answers are when the *wh*-word is interpreted either as an indefinite pronoun, or as a true *wh*-word with scope over only the embedded clause (yielding an indirect question). For this reason, direct comparison of the acceptability of yes/no answers between a declarative clause and an interrogative clause may not be very meaningful with regard to the issue of island effects in Korean. Thus, the evaluation of island effects in Korean will be primarily based on the results of the *wh*-answers here.

Overall, the results of native and heritage speakers were similar in that both groups showed *wh*-island effects in Experiments 1 and 3, but no adjunct island effects in Experiments 2 and 4. In both Experiments 1 and 3, the condition in which the *wh*-word was within the embedded *wh*-clause was noticeably worse than other conditions, indicating *wh*-island effects, which was shown by a significant interaction between the two factors, Location of *wh*-word (matrix or embedded clause) and Structure of embedded clause type (non-island or island). For heritage speakers, the effect was only marginal in Experiment 1, but it reached significance in Experiment 3. For native speakers also, the effect was smaller in Experiment 1 (DD = 0.28) than in Experiment 3 (DD = 0.71), though it was significant in both experiments. On the other hand, in Experiments 2 and 4, the acceptablity of the island-violating condition (i.e., Embedded *wh*-word inside the adjunct clause) was similar to its counterpart with the embedded *that*-clause, and no significant island effect was found. Native speakers in Experiment 2 did show an

TABLE 1 | Results summary of *Wh*-answer in Experiments 1–4.

Group		Native controls	Heritage speakers
N		48	28
Experiment 1 Canonically ordered Interrogative clause	Structure	✓ / ✓	✓ / ✓
	Location	✓ / ✓	✓ / ✓
	Interaction	✓ / ✓	# / #
	z-score (island condition)	0.19	0.22
	Raw score (island condition)	4.36 (1.48)	4.52 (1.31)
	DD score	0.28	0.23
Experiment 2 Canonically ordered Adjunct clause	Structure	✗ / ✗	✗ / ✗
	Location	✗ / ✗	✗ / ✗
	Interaction	✓ / ✓	✗ / ✗
	z-score (island condition)	0.66	0.77
	Raw score (island condition)	5.43 (1.44)	6.02 (0.89)
	DD score	−0.28	−0.13
N		48	19
Experiment 3 Scrambled Interrogative clause	Structure	✓ / ✓	✓ / ✓
	Location	✓ / ✓	✓ / ✓
	Interaction	✓ / ✓	✓ / ✓
	z-score (island condition)	−0.37	−0.06
	Raw score (island condition)	2.95 (1.38)	3.21 (1.65)
	DD score	0.71	0.72
Experiment 4 Scrambled Adjunct clause	Structure	✗ / ✗	✗ / ✗
	Location	✓ / ✓	✓ / ✓
	Interaction	✗ / ✗	✗ / ✗
	z-score (island condition)	0.42	0.55
	Raw score (island condition)	4.78 (1.55)	5.24 (1.45)
	DD score	−0.06	−0.09

✓ means "significant" (p < 0.05), # means "marginal" (p < 0.1), ✗ means "insignificant" (p > 0.1), by-subject analysis on the left, by-item analysis on the right.

interaction between Location and Structure, but in the opposite direction of what would be expected for an island effect.

CONCLUSION

Two very clear conclusions emerge from the results that we have seen in the experiments just presented. First, *wh*-clauses and adjunct clauses appear to behave very differently in Korean: *wh*-clauses behave like islands (i.e., *wh*-words within them may not take scope outside of that clause), while adjunct clauses do not (i.e., *wh*-words within them are easily able to take scope outside of that clause). This result is important in itself, because as we saw in Section Island Effects in Korean, there has been considerable uncertainty in the literature about the status of *wh*-islands in Korean.

Second, heritage speakers of Korean show essentially the same island behavior as the native controls. This was especially true in Experiments 3 and 4, where the embedded clause was scrambled and the results between the two groups were virtually identical, but even in Experiments 1 and 2, where the embedded clause was not scrambled, the two groups' results are very similar. Both heritage and native speakers thus appear to treat *wh*-clauses as islands and adjunct clauses as non-islands.

This second conclusion is particularly striking for a number of reasons. As we saw earlier, the learning environment for native and heritage speakers can be very different and this often leads to very clear language differences. Heritage speakers presumably have less overall exposure to Korean, and what exposure they have may be more limited in scope (e.g., coming from only a few speakers, rather than an entire community). In addition, heritage speakers' acquisition of Korean may have been incomplete and they may also have undergone attrition in the years since childhood. Beyond these factors relating to Korean itself, heritage speakers are also likely to be susceptible to transfer effects from English, their dominant language. It is relevant to note here that the island facts of English are different (*wh*-clauses and adjunct clauses are both islands in English), and in separate work, we have shown that these heritage speakers have native-like sensitivity to islands in English as well (Kim, 2015).

For all of these reasons, one would very reasonably expect that native and heritage speakers would differ with regard to island behavior in Korean, as they do for many other types of linguistic phenomena, but as we have seen, this is not the case. This result is consistent with the view that island phenomena are not learned, but rather follow from constraints on the way that the processor and/or grammar operates. What specific constraints could result in the type of island phenomena that we observe here for Korean? We do not offer a definitive solution here, but we do note that some very plausible possibilities have been proposed in the literature. In terms of processing, for instance, it has been claimed that a *wh*-word and a question marker need to form a dependency in *wh-in-situ* languages that is similar to the more familiar filler-gap dependency in *wh*-movement languages, and that this dependency needs to be completed as soon as possible (e.g., Miyamoto and Takahashi, 2002; Aoshima et al., 2003; Ueno and Kluender, 2009; Sprouse et al., 2011). If this dependency determines the scope of the *wh*-word, it then follows that in *wh*-clauses, scope will always be limited to that clause, since the search for a question marker will always be satisfied within that clause. This would result in the *wh*-island effect seen in Experiments 1 and 3. In adjunct clauses, on the other hand, there is no such question marker and the search continues until it is resolved outside of the adjunct clause. This leads to the lack of an island effect with adjunct clauses, as seen in Experiments 2 and 4. Alternatively, it could be that this dependency between the *wh*-word and the question marker is determined by the grammar and constrained by locality restrictions on it, as in Shimoyama's (2006) proposal for Japanese, in which the *wh*-word must associate with the question marker that is structurally closer. In this case too, though, the asymmetry in island behavior between *wh*-clauses and adjunct clauses results from the presence of a question marker in the former, but not in the latter.

Specifics aside, both of these approaches suggest that the (non-)island status of *wh*-clauses and adjunct clauses in Korean follows from fundamental properties of how the processor or the grammar operates. That is, *wh*-island violations are not possible in Korean because doing this would require a processing/grammatical operation beyond the capabilities of speakers. If this is correct, then the similarities that we have seen here between native and heritage speakers of Korean are not surprising. If heritage speakers were to not show

native-like *wh*-island effects, this would suggest that they are somehow able to surpass the processing and/or grammatical capabilities of native speakers, which hardly seems plausible. With adjunct clauses, in contrast, nothing prevents either the native or the heritage speakers from computing wide-scope readings for the *wh*-word, so neither group shows adjunct island effects. The results that we have obtained are thus exactly what is predicted by approaches in which island behavior is simply the consequence of deeper processing/grammatical traits.

If, on the other hand, island phenomena did not follow from fundamental properties of the processor and/or grammar but instead were learned from the environment, we would not predict that native and heritage island behavior would necessarily be the same. They could be, of course, but given the many differences discussed earlier in the learning environment and the possibility of transfer, it seems likely that some differences in island behavior would emerge. Since this is not what was found, our results do not lend support to this approach.

REFERENCES

Anderson, R. T. (1999). Loss of gender agreement in L1 attrition: preliminary results. *Bilin. Res. J.* 23, 389–408.

Aoshima, S., Phillips, C., and Weinberg, A. (2003). Processing of Japanese wh-scrambling constructions. *Jpn. Korean Linguist.* 12, 179–191.

Choi, Y. S. (2006). Asymmetry of locality and intervention: what really matters. *Linguist. Assoc. Korea J.* 14, 113–136.

Chomsky, N. (1971). *Conditions on Transformations*. Indiana Univ. Linguistics Club.

Culicover, P. W., and Jackendoff, R. (2005). *Simpler Syntax*. Oxford: Oxford University Press.

de Groot, C. (2005). "The grammars of Hungarian outside Hungary from a linguistic-typological perspective," in *Hungarian Language Contact Outside Hungary*, ed A. Fenyvesi (Amsterdam: John Benjamins), 351–370. doi: 10.1075/impact.20.18gro

Han, H. S. (1992). Notes on reflexive movement. *J. East Asian Linguist.* 1, 215–218.

Hofmeister, P., and Sag, I. A. (2010). Cognitive constraints and island effects. *Language* 86, 366. doi: 10.1353/lan.0.0223

Hong, S. H. (2004). On the lack of syntactic effects in Korean *wh*-Questions. *Linguist. Assoc. Korea J.* 12, 43–57.

Hwang, H. J. (2007). "*Wh*-phrase questions and prosody in Korean," in *The Proceeding of the 17th Japanese/Korean Linguistic Conference* (Los Angeles, CA: UCLA).

Ishihara, S. (2002). Invisible but Audible *Wh*-scope Marking: *Wh*-constructions and Deaccenting in Japanese. *WCCCFL* 21, 180–193.

Kim, B. (2015). *Sensitivity to Islands in Korean-English Bilinguals*. Ph.D dissertation. University of California, San Diego.

Kim, S.-W. (1989). *Wh*-Phrases in Korean and Japanese are QPs. *MIT Work. Pap. Linguist.* 11, 119–138.

Kitagawa, Y., and Hirose, Y. (2012). Appeals to prosody in Japanese *Wh*-interrogatives—Speakers' versus listeners' strategies. *Lingua* 122, 608–641. doi: 10.1016/j.lingua.2012.01.005

Kitagawa, Y. (2005). Prosody, syntax and pragmatics of *wh*-questions in Japanese. *Engl. Linguist.* 22, 302–346. doi: 10.9793/elsj1984.22.302

Kluender, R. (1998). On the distinction between strong and weak islands: a processing perspective. *Syntax Semant.* 29, 241–280.

Lee, H.-S. (1982). *Asymmetry in Island Constraints in Korean*. Unpublished manuscript, University of California Los Angeles.

Miyamoto, E., and Takahashi, S. (2002). Antecedent reactivation in the processing of scrambling in Japanese. *MIT Work. Pap. Linguist.* 43, 127–142.

Montrul, S. (2002). Incomplete acquisition and attrition of Spanish tense/aspect distinctions in adult bilinguals. *Bilingualism* 5, 39–68. doi: 10.1017/S1366728902000135

Nishigauchi, T. (1990). *Quantification in the Theory of Grammar*. Dordrecht: Kluwer.

Pearl, L., and Sprouse, J. (2013). Syntactic islands and learning biases: combining experimental syntax and computational modeling to investigate the language acquisition problem. *Lang. Acquisit.* 20, 23–68. doi: 10.1080/10489223.2012.738742

Pesetsky, D. (1987). "*Wh-in-situ*: movement and unselective binding," in *The Representation of Indefiniteness*, eds E. Reauland and A. G. B. ter Meulen (Cambridge: MIT Press), 98–129.

Polinsky, M. (2011a). Reanalysis in adult heritage language. *Stud. Second Lang. Acquisit.* 33, 305–328. doi: 10.1017/S027226311000077X

Polinsky, M. (2011b). "Annotated bibliography of research in heritage languages," in *Oxford Bibliographies in Linguistics* (Oxford University Press). doi: 10.1093/obo/9780199772810-0067

Rizzi, L. (2013). Locality. *Lingua* 130, 169–186. doi: 10.1016/j.lingua.2012.12.002

Shimoyama, J. (2006). Indeterminate phrase quantification in Japanese. *Nat. Lang. Semant.* 14, 139–173. doi: 10.1007/s11050-006-0001-5

Sorace, A. (2004). Native language attrition and developmental instability at the syntax-discourse interface: data, interpretations and methods. *Bilingualism* 7, 143–145. doi: 10.1017/S1366728904001543

Sprouse, J., Fukuda, S., Ono, H., and Kluender, R. (2011). Reverse Island Effects and the Backward Search for a Licensor in Multiple *Wh*-Questions. *Syntax* 14, 179–203. doi: 10.1111/j.1467-9612.2011.00153.x

Suh, C.-M. (1987). *A Study of the Interrogative Sentences in Korean (in Korean)*. Seoul: Top Press.

Ueno, M., and Kluender, R. (2009). On the processing of Japanese *wh*-questions: an ERP study. *Brain Res.* 1290, 63–90. doi: 10.1016/j.brainres.2009.05.084

Umeda, M. (2008). *Second Language Acquisition of Japanese Wh-constructions*. Ph.D dissertation. McGill University.

Watanabe, A. (1992). Subjacency and S-structure movement of *wh-in-situ*. *J. East Asian Linguist.* 1, 255–292.

AUTHOR CONTRIBUTIONS

BK performed the detailed design and implementation of the experiments. BK and GG both participated in the overall conception, analysis, and interpretation of the experiments, and in the drafting and revision of this article.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpsyg.2016.00134

Looking at the evidence in visual world: eye-movements reveal how bilingual and monolingual Turkish speakers process grammatical evidentiality

Seçkin Arslan[1], Roelien Bastiaanse[2] and Claudia Felser[3]*

[1] International Doctorate for Experimental Approaches to Language and Brain, University of Groningen, Groningen, Netherlands, [2] Research Group Neurolinguistics, Center for Language and Cognition Groningen (CLCG), University of Groningen, Groningen, Netherlands, [3] Potsdam Research Institute for Multilingualism, University of Potsdam, Potsdam, Germany

This study presents pioneering data on how adult early bilinguals (heritage speakers) and late bilingual speakers of Turkish and German process grammatical evidentiality in a visual world setting in comparison to monolingual speakers of Turkish. Turkish marks evidentiality, the linguistic reference to information source, through inflectional affixes signaling either direct (-DI) or indirect (-mIş) evidentiality. We conducted an eye-tracking-during-listening experiment where participants were given access to visual 'evidence' supporting the use of either a direct or indirect evidential form. The behavioral results indicate that the monolingual Turkish speakers comprehended direct and indirect evidential scenarios equally well. In contrast, both late and early bilinguals were less accurate and slower to respond to direct than to indirect evidentials. The behavioral results were also reflected in the proportions of looks data. That is, both late and early bilinguals fixated less frequently on the target picture in the direct than in the indirect evidential condition while the monolinguals showed no difference between these conditions. Taken together, our results indicate reduced sensitivity to the semantic and pragmatic function of direct evidential forms in both late and early bilingual speakers, suggesting a simplification of the Turkish evidentiality system in Turkish heritage grammars. We discuss our findings with regard to theories of incomplete acquisition and first language attrition.

Keywords: evidentiality, information source, inference, witnessing, visual world paradigm, eye-movements, Turkish-German bilingualism

Introduction

Evidentiality refers to the linguistic encoding of the type of information source an event description is based on, such as whether or not the event has been witnessed directly by the speaker (Aikhenvald, 2004). Most languages express evidentiality through lexical adverbs (e.g., *reportedly*). However, in Turkish, evidentiality is conveyed through verb inflections requiring the speaker to distinguish whether an event has been directly witnessed or has been indirectly inferred or reported

(Slobin and Aksu, 1982). In this study, we provide pioneering data on how grammatical evidentiality is processed by adult Turkish monolinguals, early bilinguals (i.e., heritage speakers of Turkish), and late bilinguals (i.e., L2 learners of German) in an eye-tracking-during-listening experiment.

Effects of bilingualism on one's native language are subject to a number of variables; in the current study, we will focus on the onset of bilingualism. Two types of bilinguals are of interest in this respect: early bilinguals (heritage speakers of a minority language) and bilingual individuals who learnt the dominant majority language after childhood. A possible consequence of bilingualism is the selective loss of properties of an individual's first language. Verbal morphology and certain syntactic constraints have been shown to be susceptible to selective erosion ('attrition') after full acquisition of the first language (De Bot and Weltens, 1991; Seliger and Vago, 1991; Yağmur, 1997; Cook, 2003; Gürel, 2004; Köpke and Schmid, 2004; Pavlenko, 2004; Köpke et al., 2007; Sorace and Serratrice, 2009). First language attrition has specifically been associated with late bilingualism. In early bilinguals (in particular, 'heritage speakers'), properties of the first language have instead been argued to be prone to disrupted acquisition processes during childhood (e.g., Montrul, 2002, 2008, 2009; Polinsky, 2006; Albirini et al., 2011, 2013). That is, early bilinguals are often assumed to not have reached full acquisition of several properties of the heritage language, due to reduced input conditions.

Köpke (2004) defines attrition as the "loss of the structural aspects of the language, i.e., change or reduction in form". In bilingual acquisition contexts, first language attrition is a possible outcome in bilinguals who acquired their second language later in life (e.g., after puberty), and after fully acquiring their first language during childhood (De Bot and Weltens, 1991; Seliger and Vago, 1991; Yağmur, 1997; Cook, 2003; Gürel, 2004; Köpke, 2004; Pavlenko, 2004; Tsimpli et al., 2004; Köpke et al., 2007). In contrast to language attrition in late bilinguals, Montrul (2002, 2008) and Polinsky (2006) have shown that an early onset of bilingualism may lead to incomplete acquisition, that is, to a failure in acquiring part(s) of the first language grammar during early childhood. Incomplete acquisition has mainly been observed in heritage speakers, who during childhood were exposed to their first language within a minority population away from where that language is spoken natively. Studies on heritage speakers of Spanish (Montrul, 2002, 2008, 2009), Russian (Polinsky, 2006, 2008), and Arabic (Albirini et al., 2011, 2013) have confirmed that several aspects of the first language grammar are subject to divergent performance and/or competence from monolingual speakers.

Montrul (2002, 2008) suggests that a disrupted acquisition process may result in unsuccessful ultimate attainment of the inherited (first) language in early bilingual adults, and that the effects of incomplete acquisition may be more severe compared to the effects of first language attrition in late bilinguals. Incomplete acquisition does not seem to affect all areas of inflectional morphology equally, however. Montrul (2009), for example, investigated adult Spanish heritage speakers' sensitivity to aspectual (preterit – imperfect) and modal (subjunctive –

indicative) distinctions using an elicited oral production task, a written morphology recognition task, and a judgment task. She found that the heritage speakers' knowledge of aspectual distinctions was better retained than their knowledge of modal distinctions, suggesting that the heritage speakers were affected by incomplete acquisition of Mood. Given that Aspect tends to be acquired earlier than Mood, Montrul (2009) attributes the heritage speakers' greater problems with Mood to maturational factors (i.e., the order of acquisition of inflectional distinctions).

Montrul's (2009) observation of Mood distinctions being eroded more than aspectual ones in Spanish heritage language is consistent with Jakobson's (1941) Regression Hypothesis, which holds that linguistic properties that are acquired late will be lost first (see Keijzer, 2010). Montrul's findings are also compatible with the Interface Hypothesis (Sorace, 2000; Sorace and Filiaci, 2006; Sorace and Serratrice, 2009), according to which linguistic properties at 'interfaces' (e.g., syntax–discourse interface) may prove particularly problematic in bilingual acquisition. Linking syntactic and discourse-level information is claimed to be particularly difficult. Sorace and Serratrice (2009) argue that "bilinguals may have fewer processing resources available and may therefore be less efficient at integrating multiple types of information in on-line comprehension and production at the syntax – pragmatics interface." Therefore, even highly proficient bilinguals may show difficulty using or processing grammatical forms that are marked in the sense of requiring very specific pragmatic licensing conditions. Sorace (2011), however, cautions against extending the Interface Hypothesis, which originally sought to account for non-target like performance patterns in near-native second language speakers, to heritage speakers.

In past few years, there has been increasing interest in understanding the properties of subtractive bilingualism, when the first language is a minority language. Most previous studies have focused on early and late bilinguals (i.e., heritage speakers and L2 speakers) living in the U. S. The nature of language erosion in bilingual individuals living in Western Europe is less well understood. Turkish is one of the most widely spoken minority languages in Germany, and it differs typologically from most of the previously studied heritage languages. Turkish is an agglutinative language with rich inflectional morphology, including the grammatical expression of evidential distinctions. The linguistic features of Turkish evidentials are described in more detail below, as well as previous experimental studies on this phenomenon.

Evidentiality in Turkish

Evidentiality refers to the linguistic encoding of a particular type of evidence for a speaker's utterance (Chafe and Nichols, 1986; Willett, 1988; Lazard, 2001; Plungian, 2001; Aikhenvald, 2004). The nature of the evidence relates to how a speaker has access to the information in his or her statement: witnessing, inference, or hearsay. Turkish expresses evidentiality through a verbal inflection paradigm with two choices for direct (witnessing) and indirect evidence (inference or hearsay), as illustrated in (1) and (2), respectively.

(1) Adam elmayı   ye**di**
    man   apple_ACC   eat _DIRECT EVIDENTIAL_
    'The man ate the apple' [witnessed]

(2) Adam elmayı   ye**miş**
    man   apple_ACC   eat _INDIRECT EVIDENTIAL_
    'The man ate the apple' [reported or inferred]

The direct evidential suffix –DI is used to refer to past events that were directly witnessed, or participated in, by the speaker. For example, in (1) –DI signals that the speaker has witnessed the apple being eaten. The indirect evidential suffixes –mIş and –(I)mIş are appropriate for use in inference or reportative contexts, respectively. For instance, in (2) the speaker has been either told that the man ate the apple, or has (physical) evidence leading him or her to infer that the man ate the apple, such as seeing peelings and leftovers of an apple on the table.

In inference contexts, the use of an indirect evidential signals non-witnessed past events that are perceived through present states or results on the basis of physical or visual evidence (Aksu-Koç and Slobin, 1986). In reportative contexts it conveys that the information is known through 'hearsay' or verbal report from a third party (Slobin and Aksu, 1982). These semantic and formal distinctions in Turkish evidentials are well understood. Several studies have indicated that the indirect evidential is the marked term on the basis of its semantic complexity since it refers to different information sources (i.e., inference and report), whilst the direct evidential is the unmarked form for referring to witnessed past events (Slobin and Aksu, 1982; Aksu-Koç and Slobin, 1986; Aksu-Koç, 1988, 2000; Sezer, 2001; Johanson, 2006). These authors also agree that while the indirect evidential bears epistemically modal connotations, the direct evidential is a non-modal term.

The use of evidentials in interrogative contexts has not been explored much in Turkish linguistics. Aikhenvald (2004) claims that evidentials in an interrogative clause reflect the type of information source available to the questioner or to the addressee. This indicates that the semantic and pragmatic uses of evidentials differ in declarative and interrogative contexts. In *wh*-interrogative clauses such as (3) and (4) below, for example, the use of a particular evidential reflects the type of information source available to the addressee of the question, while the questioner may not necessarily have access to the same information source.

(3) Hangi adam elmayı   ye**di**?
    which man   apple_ACC   eat_DIRECT EVID
    'Which man ate the apple?'

(4) Hangi adam elmayı   ye**miş**?
    which man   apple_ACC   eat_INDIRECT EVID
    'Which man ate the apple?'

The questioner's choice of a particular evidential form indicates that he or she is making assumptions on the information source available to the addressee. In (3), the questioner assumes that the addressee has witnessed who has eaten the apple; thus, a direct evidential is used. In (4), by contrast, the questioner presumes that the addressee has access to information about the event through an indirect source (e.g., inference or hearsay),

hence, an indirect evidential is used. Therefore, a particular evidential is selected in an interrogative clause depending on what the questioner assumes as to how the addressee may have acquired knowledge of the event concerned.

Experimental Studies on Turkish Evidentials

Experimental studies on evidentiality in mono- and bilingual Turkish speakers are scarce. The psycholinguistic understanding of grammatical evidentiality is limited to developmental studies in monolingual children and a small number of studies on adult bilinguals. One of the earliest empirical studies was conducted by Aksu-Koç (1988), who examined the production and comprehension of evidential morphology (among other morphemes) in Turkish-speaking children (aged 3–6). She found that the direct evidential morpheme was one of the first to be acquired, followed by the indirect evidential morpheme after a delay of about few months. Aksu-Koç (1988) notes, however, that children's early use of evidential morphemes tends to be limited to directly perceived events or present states, and that at this developmental stage children may not yet be able to distinguish the direct vs. indirect information contrast. This was confirmed by more recent studies. Öztürk and Papafragou (2007), for example, studied young monolingual Turkish children (aged 3–6) using elicited production and semantic and pragmatic comprehension tasks. The children used evidential forms appropriately but tended to have difficulty distinguishing the semantic and pragmatic content signaled by these forms. In a later study, Öztürk and Papafragou (2008) examined Turkish children (aged 5–7) using both an elicited production and a non-linguistic source monitoring task. The data reveal that Turkish children in all age groups are able to produce direct evidential forms almost faultlessly while their use of indirect evidential develops with age. Inferred and reported information sources proved more difficult for children than directly witnessed information sources even in the oldest age group; see also Ünal and Papafragou (2013). Aksu-Koç (1988) reports that monolingual Turkish children tend to gain control over the semantic and pragmatic content of direct evidentials around the age of three. The inferential readings related to the indirect evidential, however, only stabilize around the age of four in monolingual children, while reportative contexts develop around the age of four and a half. Aksu-Koç et al. (2014) and Aksu-Koç (2014) argue that modal distinctions (including epistemic readings associated with indirect evidentials) are acquired later, and that children at earlier stages of development produce non-modalized markers instead, such as the direct evidential.

Some recent studies show that evidentiality is susceptible to erosion or incomplete acquisition in Turkish heritage speakers. Arslan et al. (submitted) studied Turkish/Dutch early bilingual (i.e., second-generation heritage speakers) and Turkish monolingual adults using a sentence-verification task where participants listened to sentences containing evidential verb forms that mismatched the information contexts. For instance, an indirect evidential was mismatched to 'seen' information contexts (*Yerken gördüm, az önce adam yemeği yemiş* 'I saw the man eating; he ate_INDIRECT EVIDENTIAL the food') and a

direct evidential was mismatched to 'heard/indirect' information contexts (*Yerken görmüsler, az önce adam yemeği yedi* 'They saw the man eating; he ate_DIRECT EVIDENTIAL the food'). Participants' sensitivity to evidential verb forms was measured by asking them to press a button when a sentence was incongruent. Arslan et al. (submitted) demonstrated that the bilinguals were largely insensitive to both types of evidential mismatches. Interestingly, however, the bilinguals retained their sensitivity to tense violations (i.e., violations by past and future participles without evidentiality marked). Arslan et al.'s (submitted) data showed that evidentiality is a particularly vulnerable part of Turkish grammar in early bilingual speakers.

Furthermore, Arslan and Bastiaanse (2014) investigated narrative speech production in second-generation Turkish/Dutch early bilingual adults. The early bilinguals made a large number of substitution errors by inappropriately using direct evidentials in contexts that required an indirect evidential form. The early bilingual adults showed reduced sensitivity to the semantic distinctions between information sources that the evidential forms signal. Arslan and Bastiaanse (2014), nonetheless, report that the early bilingual adults did not substitute the indirect evidential where a direct one should be produced. The authors suggest that the indirect information source is incorporated while direct evidence is ignored, as if the direct evidential does not carry an evidential value in early bilingual Turkish speakers' oral production.

Summarizing, previous studies indicate (i) that the direct evidential is acquired earlier than the indirect evidential, possibly due to the latter being more complex in terms of its semantics (e.g., Aksu-Koç, 1988; Öztürk and Papafragou, 2007, 2008); (ii) that evidential terms in Turkish are highly susceptible to erosion in adult heritage speakers (Arslan and Bastiaanse, 2014; Arslan et al., submitted). The studies discussed above have also left some questions unexplored. First, it is not clear whether insensitivity to evidentiality distinctions is restricted to early bilingual heritage speakers or whether it can also be observed in late bilinguals. Second, although Arslan et al. (submitted) measured the processing of evidentiality using a response-time task, the moment-by-moment time course of processing evidentiality has not been investigated yet. Finally, recall that the use of evidential forms is linked to the kind of evidence available to the speaker (in declarative clauses) or the addressee (in interrogative clauses), and nothing is known as yet about how comprehenders interact with this evidence during their processing of grammatical evidentiality.

In the current study, we carried out an eye-movement monitoring experiment with three groups of participants: early and late Turkish/German bilinguals and a reference group of monolingual Turkish speakers. Testing two different bilingual groups should allow us to explore whether differences in the age of bilingualism onset affects bilinguals' processing of evidentiality. The aim of the experiment was to unveil the nature of processing evidentiality through monitoring participants' eye movements while they listened to sentences with grammatical evidentiality in a visual-world paradigm. This is a very compelling way to test processing of evidentiality as the visual-world paradigm allows us to measure participants'

moment-by-moment eye-movements while they interact with different types of visual evidence. Our visual stimuli included picture pairs that encoded either 'witnessed' or 'inferable non-witnessed' events, which were appropriate for the use of direct and indirect evidential forms, respectively. In particular, we sought to answer the following questions:

- Do early and late bilinguals differ from monolinguals in their processing of evidentiality?
- Do monolingual, early and/or late bilingual Turkish speakers differ in their processing of direct and indirect evidentials?

Given the findings of previous studies on early bilingual heritage speakers living in the U. S., inflectional morphology seems to be particularly affected. This is consistent with Arslan et al.'s (submitted) findings for early bilingual speakers of Turkish in the Netherlands. Considering these data, we expect early bilinguals to show a reduced sensitivity to evidentiality in comparison to monolingual Turkish speakers. If this is a consequence of incomplete acquisition, then early bilinguals will also be sensitive to evidentiality compared to late bilinguals, who we expect to pair with the monolinguals. The hypotheses we introduced above moreover predict an asymmetrical insensitivity in bilingual participants' responses to direct and indirect evidential forms. Specifically, the Interface Hypothesis predicts more problems during bilinguals' processing of the indirect than the direct evidential forms. According to this hypothesis, integrating information from multiple linguistic domains – in particular, integrating morphosyntactic and pragmatic information – is difficult for speakers who have not fully acquired the language under investigation. Recall that the use of indirect evidentials is licensed only in specific pragmatic contexts that require more or less complex inferential reasoning, whereas direct evidentials are used as an 'elsewhere' form in the absence of such contexts, signaling that an event was witnessed directly. The Regression Hypothesis also predicts more problems in bilinguals' responses to indirect than to direct evidential forms as the former are acquired later in development.

Materials and Methods

Participants

Sixty-one adult Turkish speakers were recruited from the Turkish community of Berlin, Germany. They were categorized into three groups on the basis of their age of onset of bilingualism. A group of early bilinguals ($n = 19$), who were all born in Germany (i.e., second generation heritage speakers of Turkish), and a group of late bilinguals ($n = 20$) were recruited. The late bilinguals were L2 learners of German who came to Berlin after puberty (i.e., after the age of 13). Finally, a reference group of monolingual Turkish speakers ($n = 22$) who had no previous contact with German also participated. A demographic information questionnaire was completed by all participants. In addition, the bilinguals responded to a short language test in both German and Turkish, adapted from the Goethe (Goethe-Institut e.V.) and telc (telc GmbH) placement tests; see **Table 1**.

TABLE 1 | Numbers and age of participants, AoA = age of acquisition in years with min-max age range, and proficiency test scores (ranges in brackets) in Turkish and German for bilingual participants.

	N	Age	AoA Turkish	AoA German	Turkish score	German score
Monolingual	22	24 (20–36)	From birth	NA	NA	NA
Late bilingual	20	30 (21–46)	From birth	13–27	89.5% (63–100)	61.3% (23–93)
Early bilingual	19	27 (22–36)	From birth	1–4	71.1% (13–100)	91.2% (76–100)

The monolinguals were native Turkish speakers from Turkey who were in Berlin for holidays or family visits during the time they were recruited. None of them spoke any German. All participants were highly educated (i.e., college students or graduates) and spoke the standard Turkish dialect. No speakers of any ethnical languages or dialects participated in this study. The participants had normal hearing and (corrected to normal) vision. They gave their consent under the Helsinki declaration and were paid a fee of 10 Euros.

Materials

Sixty visual displays, each comprising a pair of photos presented next to each other, were created as shown in **Figure 1**. One of the photos was the target picture and the other one served as a context picture. To create the visual displays, 20 action verbs were combined with six different people and 10 different inanimate objects (i.e., *süt içmek* 'to drink milk'). The same actions were displayed in two experimental conditions, a direct and an indirect evidential one, as well as in a non-evidential distractor condition involving the future tense (*n* = 20 each). The photographs used in this experiment were taken from European, Asian, and African versions of the Test for Assessing Reference of Time: TART (Bastiaanse et al., 2008). Different 'models' from different versions of TART were used with the same action displayed in different conditions in a counterbalanced manner. For example, drinking milk appeared once in the direct evidential condition acted by a European-looking person, once in the indirect evidential condition acted by a person of Asian appearance, and once in the future tense condition acted by a person of African appearance as shown in **Figure 1**. An equal number of male and female 'models' appeared in each condition.

To encode direct and indirect evidentiality contexts visually, different states of the same action were represented next to each other. For the direct evidential condition, an action was shown while it was happening in one of the photographs and its end-state in the other (see **Figure 1A**). This was an example of a witnessed event, appropriate for the use of a direct evidential form. For the indirect evidential condition (**Figure 1B**), an action was displayed in its end-state and in a 'pre-action' state, that is, before the action was initiated. This means that the action could only possibly be inferred, making this kind of visual display appropriate for the use of an indirect evidential form. In both evidential conditions, the target picture was the photograph that depicted the end-state of the action. For the future tense condition (**Figure 1C**), an action was shown in the target photo in its pre-action state. The future items also included a context photo, which was showing the action as ongoing in half of the
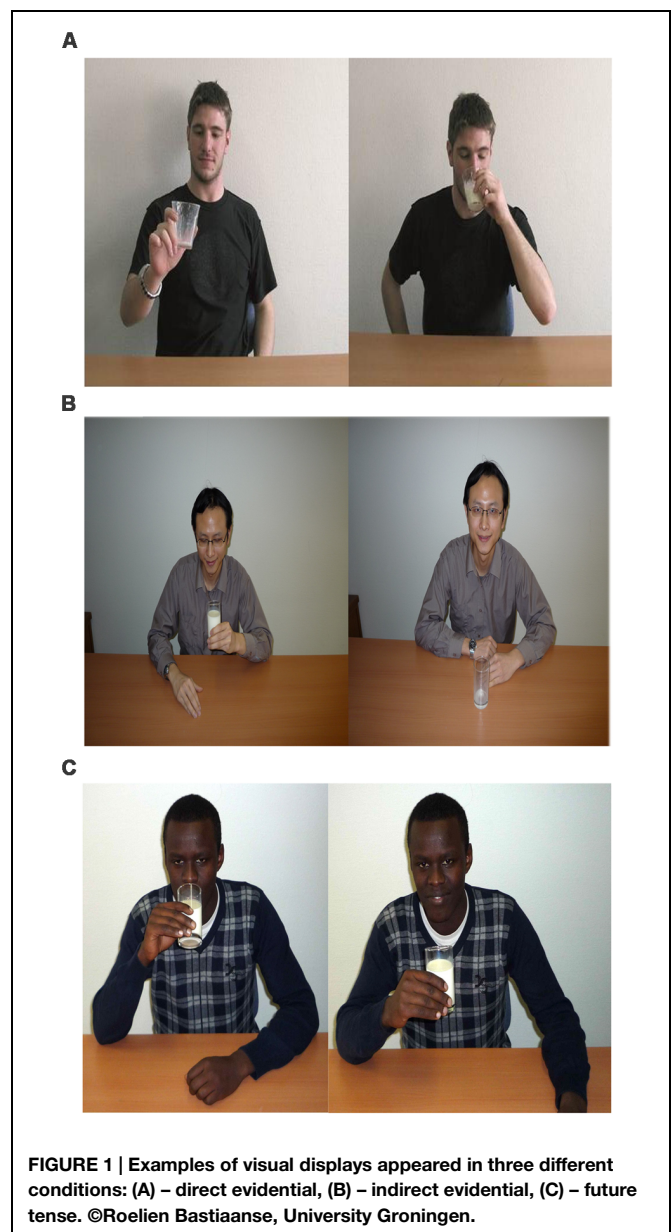
FIGURE 1 | Examples of visual displays appeared in three different conditions: (A) – direct evidential, (B) – indirect evidential, (C) – future tense. ©Roelien Bastiaanse, University Groningen.

future items, and in its end-state in the other half. The order of the two photographs was reversed in half of the items so that the target picture did not always appear on the same side.

The auditory stimuli consisted of interrogative clauses that were read by a female Turkish native speaker and digitally

recorded. Examples for each of the three conditions are given in (5)–(7) below. In the two evidential conditions, the participants were asked to identify the picture showing the result of the action. In the future tense condition, the target picture was the one depicting a pre-action state (e.g., with the glass of milk still full and untouched).

(5) Direct evidential
Hangi   fotoğraftaki   adam   dün   sütü
which   photograph_LOC man   yesterday   milk_ACC
içti   ender   bir   istekle?
drink_DIRECT EVID unusual   one   desire
'In which photograph did the man drink the milk yesterday with an unusual desire?'

(6) Indirect evidential
Hangi   fotoğraftaki   adam   dün   sütü
which   photograph_LOC man   yesterday   milk_ACC
içmiş   ender   bir   istekle?
drink_INDIRECT EVID unusual   one   desire
'In which photograph did the man drink the milk yesterday with an unusual desire?'

(7) Future tense (non-evidential)
Hangi   fotoğraftaki   adam   birazdan   sütü
which   photograph_LOC man   soon   milk_ACC
içecek   ender   bir   istekle?
drink_FUTURE unusual   one   desire
'In which photograph will the man drink the milk soon with an unusual desire?'

A three-word padding phrase (e.g., *ender bir istekele* 'with unusual desire') was added at the end of each interrogative clause to preclude the auditory stimuli from terminating at the critical verb. Extending the stimuli sentences in this way was necessary so as to extend measuring time and thus enable us to capture potential spillover effects, and to reduce the possibility of our eye-movement data being affected by global end-of-sentence wrap-up processes.

Evaluation of the Experimental Sentence Stimuli[1]

The plausibility of our experimental stimuli was evaluated in an offline rating study using a four-point Likert scale (1 = very plausible, 4 = very implausible). To construct plausible test items, the evidentiality sentences exemplified by (5) and (6) were converted into declarative clauses. The 'plausible direct evidential condition' ($n = 20$) contained semantically coherent sentences with a direct evidential form (e.g., *adam dün sütü içti, ender bir istekle* 'the man drank the milk with unusual desire'), and the 'plausible indirect evidential condition' ($n = 20$)

---

[1] A reviewer suggests that the use of evidential sentences with the padding phrases positioned at the end of the sentences sounds rather unnatural, especially for the indirect evidential sentences. The reviewer claims that the indirect evidentiality sentences used in the current study cannot be combined with adverbial phrases such as *ender bir istekele* 'with unusual desire' since the indirect evidential signals a "non-witnessed" event. This is on the assumption that in inference contexts, where there is nobody who actually witnessed how the action was performed, adverbials of this kind cannot be used to modify the action. The purpose of our offline rating task was to ascertain whether our direct and indirect evidentiality stimuli sounded equally plausible.

contained semantically coherent sentences with an indirect evidential form (e.g., *adam dün sütü içmiş, ender bir istekle* 'the man drank the milk with unusual desire'). To create implausible counterparts of the plausible conditions, the agent and theme arguments in those sentences were reversed (e.g., *süt dün adamı içti, ender bir istekle* 'the milk drank the man with unusual desire'). The plausible and implausible sentences were distributed across four presentation lists, counterbalanced across participants. Sentences constructed with a same verb in different conditions appeared in different lists so as to minimize potential effects of repetition. In addition, 30 plausible and implausible filler sentences were added to each list, resulting in a total of 50 items per list.

Participants included 43 monolingual speakers of standard Turkish (mean age = 26.3, range = 17–45, 24 males), none of whom took part in the main eye-tracking experiment. All participants were living in Turkey and none of them reported to speak any foreign language proficiently. The rating task was administered as a web-based questionnaire. At the beginning of the task, the following instructions were provided in Turkish: "You are being asked to rate the plausibility of some Turkish sentences (i.e., how 'intuitive and reasonable' do these sentences sound to you). Please read each sentence carefully and click on one of the answer choices provided under each sentence. On every page, there are five sentences. When you have finished rating the sentences on one page, click on 'continue,' and when you have finished rating all of the sentences, please click on 'submit'."

The results showed that the plausible direct evidential condition was rated significantly more favorably than its implausible counterpart [1.66 vs. 3.73, $t(42) = -19.4$, $p < 0.0001$], and the plausible indirect evidential condition was rated as more plausible than its implausible counterpart [1.60 vs. 3.83, $t(42) = -23.3$, $p < 0.0001$]. Crucially, participants' ratings of the plausible direct and indirect evidential conditions did not differ statistically [$t(42) = 1.39$, $p = 0.17$], and neither did their ratings of the two implausible conditions [$t(42) = 1.76$, $p = 0.09$].

Procedure

Presentation of visual and audio stimuli was programmed in two lists by using the SMI experiment builder software (SensoMotoric Instruments GmbH). A participant saw two photos presented next to each other in each trial, as described above. The evidential items were counterbalanced across participants over the two lists, so that an evidential item only appeared in either the direct or the indirect evidential condition. Each participant saw 10 direct and 10 indirect evidential items. In addition, 20 future tense items were added to each list as non-evidential distractor items. Therefore, each participant was exposed to an equal number of evidential and non-evidential items. A further 20 filler items, containing a subject participle complement clause (i.e., a non-finite verb form: *Hangi fotoğraftaki adam dün yemeği pişiren adam* 'which photograph_LOC man yesterday food_ACC cook SUBJECT PARTICIPLE man?'), were added so that each presentation list contained 60 items. Presentation of the

auditory stimuli was delayed by 1 s with respect to the visual stimuli in all items. Pauses were programmed after every block of 20 items. The items were presented in a randomized manner.

Participants were tested individually in a dedicated testing room in Berlin. They were asked to sit within a convenient sight distance from a 1680 pixels × 1050 pixels-wide (i.e., 22 inches) PC screen. They were then given the following instructions in Turkish: "You are about to begin an eye-tracking experiment. Please listen to the sentences carefully, and click on the photograph that corresponds to the sentences you hear. When you click the next item will begin." Two practice trials were presented during which the participants were provided with feedback and the opportunity to ask questions if they had any. Before the main eye-tracking experiment began, participants were reminded not to turn their gaze off the screen. When participants responded, the presentation of the next stimulus was initiated manually by the experimenter. Eye movements were monitored and sampled at a rate of 60 Hz, one frame per 16 ms, by a remote SMI eye-tracking system positioned underneath the stimulus screen. The research was approved by the ethics committee of the University of Potsdam (application number 37/2011).

Analysis

Three types of dependent variables were obtained and analyzed separately: accuracy of clicks, response times (RTs), and proportion of looks. The accuracy data were analyzed using generalized linear mixed-effects regression models, and the RTs data using linear mixed-effects regression models (Baayen, 2008). RTs that exceeded three standard deviations beyond the group means were excluded. Any responses made before the onset of the critical verbs were rejected (around 1.5%). For the proportions of looks analysis, a time window of 2000 ms from the onset of the critical verb was selected.[2] The first 200 ms after verb onset were excluded from this time window, since it takes about 200 ms to program and execute an eye movement (Rayner et al., 1983). Proportion of looks was a binary variable indicating whether the participants fixated on the target picture or not. We excluded 0.92% of the data due to off-screen looks. The analyses were done on non-aggregated data. Participants' proportion of looks were analyzed with mixed-effects multilevel logistic regression models (Barr, 2008), using the 'lme4' and 'multcomp' statistical packages of R version 3.1.1 (R-Core-Team, 2012).

Results

Accuracy and Response Times

Mean accuracy and RTs data are shown in **Table 2** and the fixed effects from mixed-effects regression models performed on accuracy and RTs of responses are given in **Table 3**. For

TABLE 2 | Mean proportion of accuracy, standard error rates (SE), and response times (RTs) of click responses.

	Monolingual	Late bilingual	Early bilingual
Accuracy			
Direct evidential	0.89 (0.02)	0.67 (0.03)	0.63 (0.03)
Indirect evidential	0.89 (0.02)	0.90 (0.02)	0.85 (0.03)
Future tense (distractor)	0.93 (0.01)	0.90 (0.01)	0.88 (0.02)
RTs			
Direct evidential	2214.7	2707.4	2716.5
Indirect evidential	2262.3	2339.7	2494.3

the accuracy data, significant effects of group with negative estimate values indicate that both late and early bilinguals were less accurate than monolinguals.[3,4] However, the between-groups differences were modulated by condition, as witnessed by significant interactions between the factors group and condition. Therefore, *post hoc* analyses were performed using Tukey tests. These revealed that both late ($\beta = 0.213$, $SE = 0.04003$, $z = 5.326$, $p > 0.001$) and early bilinguals ($\beta = 0.228$, $SE = 0.035$, $z = 6.418$, $p < 0.001$) responded less accurately to the direct evidential than to the indirect evidential condition, whereas the monolinguals showed no difference between the two conditions ($\beta = 0.0105$, $SE = 0.029$, $z = 0.353$, $p = 0.072$). There were group differences in participants' responses in the direct evidential condition, with both the early ($\beta = -1.897$, $SE = 0.5404$, $z = -3.511$, $p = 0.0012$) and the late bilinguals ($\beta = -1.685$, $SE = 0.5311$, $z = -3.172$, $p = 0.0042$) less accurate than the monolinguals. The early and late bilinguals did not differ in their responses in the direct evidential condition ($\beta = 0.212$, $SE = 0.5005$, $z = 0.424$, $p = 0.905$). For participants' responses in the indirect evidential condition, no within or between group differences were observed (all ps > 0.346).

With regard to RTs, the model outputs shown in **Table 3** revealed significant effects of group but not of condition. The negative estimate values of the group effects confirm that both late and early bilingual groups were slower in their responses than monolinguals irrespective of condition. Since the interactions between group and condition were also significant, *post hoc* analyses were performed. Both the late ($\beta = 372.10$, $SE = 116.10$, $z = -3.204$, $p = 0.001$) and early bilinguals ($\beta = 332.90$, $SE = 150.0$, $z = -2.22$, $p = 0.026$) showed longer RTs to the direct evidential condition than to the indirect evidential condition, whereas no significant between-condition difference was seen

---

[2]The mean onset of the critical verbs was 4162 ms after each trial began, minus 1000 ms silence, and the mean sentence offset time was as 5470 ms from the beginning of the sentences.

[3]An initial model was built with future tense items included, which showed no effects of condition for indirect evidential vs. future tense items ($\beta = -0.501$, $SE = 0.289$, $z = -1.731$, $p = 0.082$), and for direct evidential vs. future tense ($\beta = -0.528$, $SE = 0.286$, $z = -1.840$, $p = 0.065$). Effects of group were not found, as well: late bilinguals vs. monolinguals ($\beta = -0.3109$, $SE = 0.3758$, $z = -0.827$, $p = 0.40$), and for early bilinguals vs. monolinguals ($\beta = -0.4961$, $SE = 0.3752$, $z = -1.322$, $p = 0.18$). As the future items were used as distractors, they were omitted from all subsequent analyses.

[4]The accuracy of responses in the direct and indirect evidential conditions in the late bilingual group correlated with their Turkish ($r = 0.102$, $p = 0.041$) and German ($r = 0.184$, $p < 0.001$) language proficiency scores, whereas no such correlations were found in the early bilingual group (both ps > 0.36), as shown by Pearson tests.

TABLE 3 | Fixed effects from the generalized linear mixed-effects regression models performed on accuracy of clicks and linear mixed-effects regression model performed on RTs.

Fixed effect	Accuracy of clicks				RTs of clicks			
	Estimate	SE	Z-value	p-value	Estimate	SE	t-value	p-value
Intercept	2.472	0.353	6.994	<0.001***	2282.28	200.04	11.409	<0.001***
Condition (indirect evidential)	0.056	0.310	0.182	0.855	−25.08	107.34	−0.234	0.815
Group (late bilingual)	−1.737	0.407	−4.262	<0.001***	525.21	268.52	1.956	0.050*
Group (early bilingual)	−1.531	0.403	−3.798	<0.001***	554.36	274.24	2.021	0.043*
Condition × Group (late bilingual)	1.393	0.426	3.266	0.001**	−420.22	160.05	−2.626	0.009**
Condition × Group (early bilingual)	1.762	0.443	3.977	<0.001***	−349.25	169.03	−2.066	0.039*
	Formula: accuracy ~ Condition * Group + (1| subject_no) + (1| item_no)				Formula: RTs ~ Condition * group + (1| item_no) + (1| subject_no)			

*p < 0.05, **p < 0.01, ***p < 0.001.

in the monolinguals ($\beta = -29.31$, $SE = 100.30$, $z = -0.292$, $p = 0.77$). Within the responses in the direct evidential condition, group contrasts proved significant. Both the early ($\beta = -475.26$, $SE = 156.45$, $z = -3.038$, $p < 0.01$) and the late bilinguals ($\beta = -401.01$, $SE = 150.37$, $z = -2.667$, $p = 0.020$) responded slower than the monolinguals, whereas late bilinguals did not differ from the early bilinguals ($\beta = -74.25$, $SE = 168.33$, $z = -0.441$, $p = 0.77$). Within the responses in the indirect evidential condition, by contrast, no group differences were found (all $ps > 0.14$).

Proportions of Looks

Figure 2 illustrates the moment-by-moment changes in participants' proportions of looks toward the target picture for the direct and indirect evidential conditions during the entire 2000 ms time window, and **Figure 3** shows the mean proportions of looks in the main and later time windows, respectively. **Figure 2** indicates that the proportions of looks to the target picture were around 50% (i.e., participants gazed on both the target and context photographs with equal likelihood) at the beginning of the time window for all groups, which confirms that participants did not visually prefer one photograph over the other before they heard the critical verb form. As we mentioned above, any fixation changes prior to 200 ms from verb onset cannot be attributed to the critical stimulus.

Visual inspection of the eye-movement data indicated that during the initial 200–1000 ms after verb onset, both bilingual groups' eye movements tended to oscillate between the target and context pictures, and that a more stable increase in looks to the target picture only emerged after about 1000 ms (see **Figure 2**). The monolinguals, however, showed more stable eye-movement patterns, with looks to the target pictures starting to increase rather steeply from about 600 ms onwards in both the direct and the indirect evidential conditions. The monolingual group's proportion of looks to the target picture reached a peak at around 1200 ms. After 1200 ms, the monolinguals started turning their gaze to the context picture, where the actions were shown to be in progress, in the direct evidential condition. They kept fixating the target photo during the processing of indirect evidentials in the same time window. Therefore, on the basis of this visual inspection, two time windows were chosen for the statistical

analyses: (i) the 'main' time window (200–2000 ms), and (ii) a 'late' time window (1200–2000 ms); see **Figure 3**.

The fixed effects of the mixed-effects logistic regression models built on the proportion of looks data from the main and late time windows are shown in **Table 4**. Since proportion of looks data do not display a linear relationship with time, in addition to linear time, quadratic, and cubic time variables were included in the models so that fixation changes over time can be best captured.

Outcomes from the model for the main time window showed significant effects of group, with both early and late bilinguals fixating less frequently on the target picture within the main time window compared to the monolinguals. Significant interactions between condition and group were found which indicate between-group differences in participants' eye-movement patterns across the two experimental conditions.

Within the main time window, fixations on the target picture were found to be reduced in the direct evidential condition in both the early ($\beta = 0.0518$, $SE = 0.0052$, $z = 9.857$, $p < 0.0001$) and late bilinguals ($\beta = 0.0253$, $SE = 0.0051$, $z = 4.911$, $p < 0.0001$) in comparison to the number of target fixations in the indirect evidential condition. The monolingual group showed no difference between the two evidential conditions ($\beta = -0.0046$, $SE = 0.005$, $z = -0.919$, $p = 0.35$), as was confirmed by Tukey tests.

The early bilinguals fixated less on the target picture than the monolinguals in the direct evidential condition ($\beta = -0.09448$, $SE = 0.03616$, $z = -2.613$, $p = 0.024$), while the late bilinguals differed only marginally from the monolinguals here ($\beta = -0.07704$, $SE = 0.03554$, $z = -2.168$, $p = 0.077$). The late and early bilinguals did not differ from each other in the direct evidential condition ($\beta = 0.01744$, $SE = 0.03617$, $z = 0.482$, $p = 0.87$), however. For the indirect evidential condition, no between-group differences were found (all $ps > 0.67$).

For the late time window (see **Table 4**), the model outputs showed effects of condition, group, as well as interactions between these two factors. To investigate the nature of these differences, *post hoc* analyses were performed. During their processing of direct evidentials, both late ($\beta = -0.11413$, $SE = 0.04053$, $z = -2.816$, $p = 0.013$) and early bilinguals ($\beta = -0.12507$, $SE = 0.04115$, $z = -3.040$, $p = 0.006$) looked less frequently toward the target picture than the monolinguals
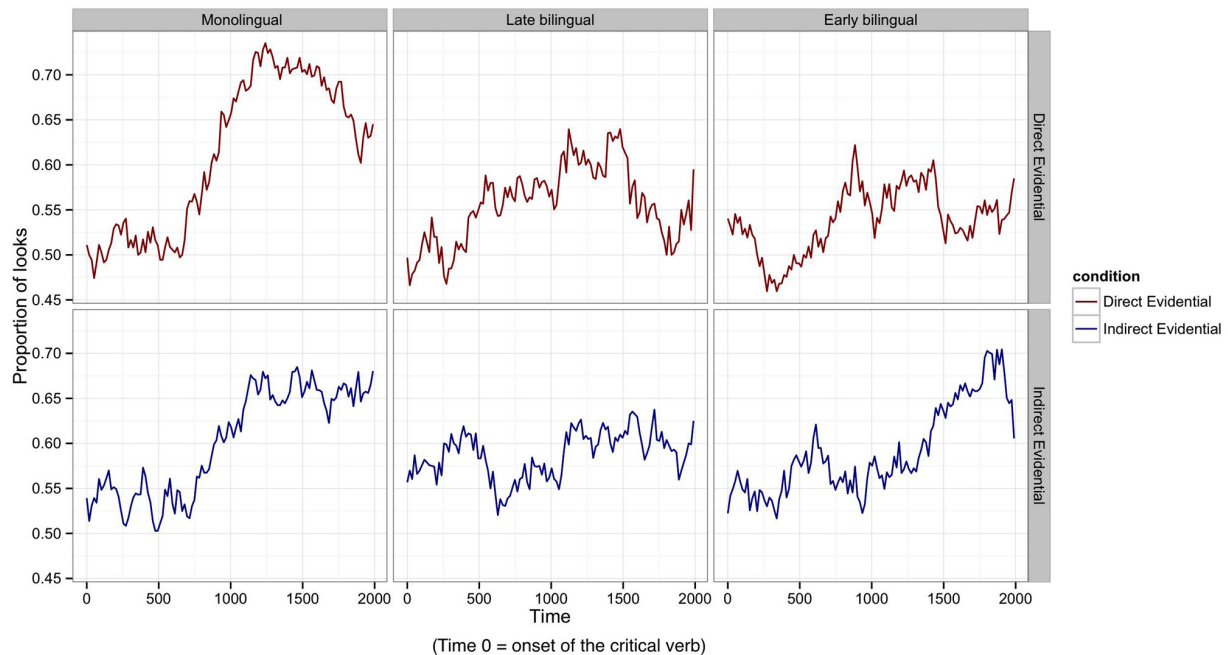
FIGURE 2 | Mean proportions of target fixations per participant group and condition for the 2000 ms time window from the onset of the critical verb. The y-axis shows participants' mean fixation proportions for each of the two evidentiality conditions.
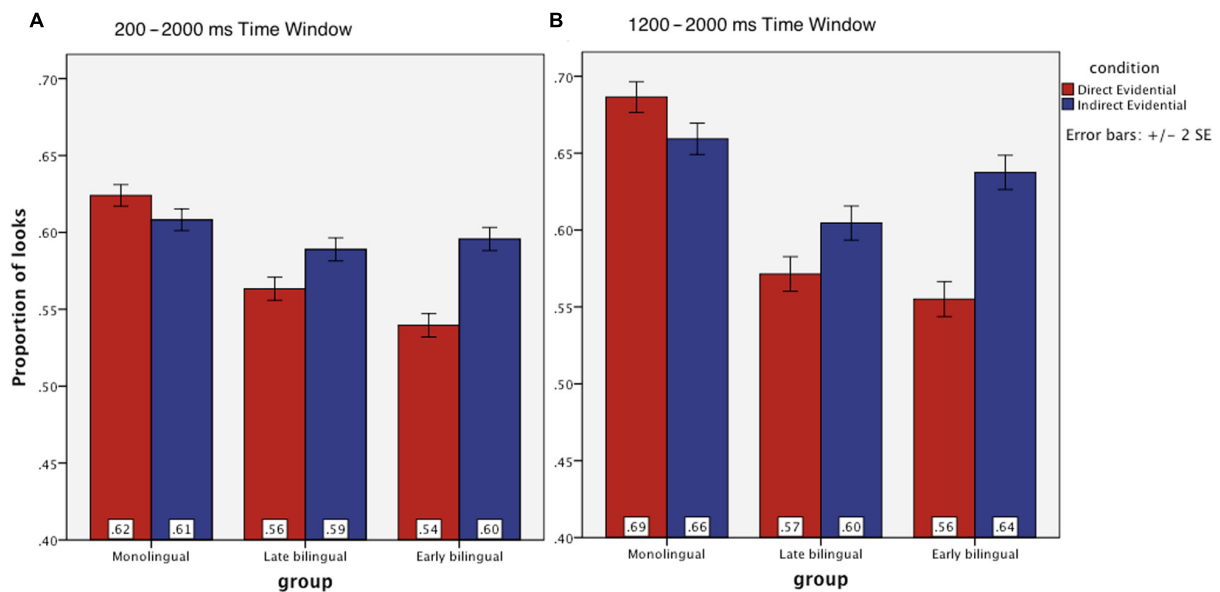
FIGURE 3 | Mean proportions of target fixations in three groups of participants during their processing of direct and indirect evidentials in two different time windows: (A) 200–2000 ms and (B) 1200–2000 ms after verb onset.

did. Again, no significant between group differences were found during participants' processing the indirect evidentials (all $ps > 0.44$).

Within-group comparisons revealed that both the early ($\beta = 0.077061$, $SE = 0.0077$, $z = 9.98$, $p < 0.0001$) and the late bilinguals ($\beta = 0.034811$, $SE = 0.0075$, $z = 4.599$, $p < 0.0001$)

fixated more frequently on the target picture in the indirect than in the direct evidential condition during the late time window. The monolinguals showed the opposite pattern: they looked at the target picture slightly more frequently in the direct than the indirect condition ($\beta = -0.015209$, $SE = 0.0072$, $z = -2.017$, $p = 0.035$).

TABLE 4 | Fixed effects from the mixed-effect logistic regression model performed on the proportion of looks data in the main time window (200–2000 ms) and late time window (1200–2000 ms).

Fixed effect	Main time window (200–2000 ms)				Late time window (1200–2000 ms)			
	Estimate	SE	t-value	p-value	Estimate	SE	t-value	p-value
Intercept	5.674	1.092	5.198	<0.001***	−4.241	8.824	−0.481	0.631
Linear time	−1.012	1.088	−0.001	0.999	2.123	7.795	0.027	0.978
Quadratic time	2.451	2.800	8.753	<0.001***	−1.320	1.053	−1.253	0.210
Cubic time	−9.140	9.202	−9.933	<0.001***	2.648	2.193	1.207	0.227
Condition (indirect evidential)	−4.365	4.764	−0.916	0.360	−2.148	7.260	−2.958	0.003**
Group (early bilingual)	−7.072	2.459	−2.876	0.004**	−1.252	3.081	−4.063	<0.001***
Group (late bilingual)	−5.331	2.426	−2.197	0.028*	−1.023	3.041	−3.364	0.001**
Condition × Group (early bilingual)	5.432	6.944	7.823	<0.001***	1.018	1.060	9.601	<0.001***
Condition × Group (late bilingual)	3.246	6.883	4.716	<0.001***	5.260	1.051	5.003	<0.001***

Formula: PropLook ~ Linear time + Quadratic time + Cubic time + Condition * Group + (1 + Linear time | participants) + (1 + Linear time | items)

$^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Notwithstanding the monolingual participants' overall higher number of fixations on the target picture in the direct evidential condition in the late time window, they tended to shift their gaze toward the context photo from about 1200 ms in the direct evidential condition whereas they kept fixating on the target photo in the indirect evidential condition (see **Figure 3**). To further examine these eye-movement changes over time, we ran the model again on the monolingual eye-movement data from the late time window with fixed effects of linear time and condition. The model output showed a significant effect of linear time ($\beta = -1.243$, $SE = 2.094$, $t = -5.937$, $p < 0.001$), condition ($\beta = -1.954$, $SE = 4.80$, $t = -4.071$, $p < 0.001$), and an interaction between the two factors ($\beta = 1.128$, $SE = 2.966$, $t = 3.804$, $p < 0.001$). These results confirm that the monolinguals' fixation changes over time within the late time window were different in the direct and indirect evidential conditions.[5]

Summary of Results

Both the late and the early bilinguals were slower and less accurate than the monolinguals in their responses in the direct evidential condition, whereas they patterned with the monolinguals in the indirect evidential condition. Furthermore, within the response data there were interactions with group, showing that both the late and early bilinguals responded less accurately to the direct than to the indirect evidential condition, while the monolinguals showed no difference between these two conditions. A similar contrast was found in response latencies.

[5]Participants' eye-movement changes over time in the late time window also showed different group characteristics within each condition. In the direct evidential condition, there were effects of linear time ($\beta = -1.053$, $SE = 1.287$, $t = -8.181$, $p < 0.001$), and of group ($\beta = -6.760$, $SE = 2.542$, $t = -2.659$, $p < 0.01$). In the indirect evidential condition, by contrast, there was an effect of linear time ($\beta = 3.844$, $SE = 1.259$, $t = 3.054$, $p < 0.001$) but not of group ($\beta = -2.731$, $SE = 13.317$, $t = -0.823$, $p = 0.41$). Eye-movements changed over time in both condition, as linear time was significant in both conditions. However, there was an effect of group in the direct evidential (but not in the indirect evidential) condition suggesting that the moment-by-moment eye-movements changes in the late time-window are different for individual groups in the direct evidential condition, but similar in the indirect evidential condition.

These behavioral results were reflected in the proportion of looks data. Bilinguals were less likely to look at the target picture in the direct compared to the indirect evidential condition in both the main and the late time windows. In the late time window (i.e., from 1200 ms onwards), the monolinguals shifted their gaze toward the context picture during their processing of direct evidentials, whilst the bilinguals' eye-movements tended to oscillate more between the target and context photos.

Discussion

The results reported add to our understanding of how evidential morphology is processed and linked to the type of evidence available by both mono- and bilingual Turkish speakers. Our first research question was whether bilinguals differ from Turkish monolinguals in processing evidentiality. The second question was whether monolingual, late and/or early bilingual Turkish speakers differ in their processing of direct vs. indirect evidentials.

The answer to the first question is clearly positive, as early and late bilinguals were found to differ from the monolinguals in their end-of-trial responses and eye-movement patterns. Both late and early bilinguals responded less accurately and looked less often to the target picture when processing direct evidentials compared to the monolinguals. Regarding our second research question, we observed an interesting asymmetry between the direct and indirect evidential conditions in the two bilingual groups that was absent in the monolingual group. Both early and late bilinguals showed greater problems processing direct compared to indirect evidentiality. This asymmetry was reflected in reduced response accuracy, longer response latencies, and in a lower proportion of looks to the target picture, in the direct compared to the indirect evidential condition. No statistical between-group differences were found for early vs. late bilinguals, indicating that the onset of bilingualism did not affect the way they processed evidentiality.

How can the observed pattern of results be accounted for? Previous studies have shown that bilinguality may affect the way people use or process their native language, with

bilinguals – in particular, heritage speakers – often performing differently from monolinguals on linguistic tasks. The age of bilingualism onset has been argued to be an important factor: whilst non-target like performance in late bilinguals is often attributed to first language attrition, non-target like performance in early bilinguals has been associated with incomplete acquisition. In first language attrition, individuals who initially acquired their native language fully may lose certain properties of that language later in life, possibly influenced by properties of a second language. In incomplete acquisition, by contrast, early bilinguals (or heritage speaker) experience disrupted acquisition processes, as a result of which certain properties of their native language are never properly acquired.

In Turkish child language acquisition, the indirect evidential is acquired after the direct evidential; it is conceivable that our early bilinguals did not fully acquire the correct use of indirect evidentials as compared to the late bilinguals. Incomplete acquisition in early bilinguals has also been associated with more severe outcomes in comparison to attrition in late bilinguals (Montrul, 2002, 2008). This is not what we found, however. Both bilingual groups were at the monolingual level in processing indirect evidentiality but performed worse than the monolinguals in the direct evidential condition. We did not find any differences between early and late bilinguals' responses in the direct evidential condition, which means that both bilingual groups were equally affected in their processing of direct evidentiality in comparison to the monolinguals. Our results, thus, do not indicate that an earlier onset to bilingualism results in more severe effects than a later onset of bilingualism.

We believe that the late bilinguals in our study were affected by a form of attrition. However, on the basis of the current data, for the early bilinguals it is impossible to precisely tease apart effects of attrition from those of incomplete acquisition. Studies on monolingual children's acquisition of evidential morphology are still scarce. These studies suggest that by the age of six, the conceptual development linked to the use of indirect evidential forms is not yet fully complete (e.g., Öztürk and Papafragou, 2007, 2008). It is thus unclear at which age the development of the evidential system finalizes. The fact that both bilingual groups showed reduced sensitivity to direct evidentials but were at the monolingual level in their processing of indirect evidentials indicates that the representation and/or pragmatic function of the direct evidential morpheme differs between mono- and bilingual Turkish speakers. This suggests that the underlying reason for the observed between-group differences is not related to the age at which the bilinguals' acquired German but to the linguistic properties of evidentiality.

Recall that Turkish indirect evidentials are assumed to have modal properties unlike direct evidentials, and that the former are thought to be semantically more complex that the latter. Turkish linguists also agree that the direct evidential is the 'unmarked' evidential form (e.g., Aksu-Koç, 1988, 2000; Sezer, 2001; Johanson, 2006), while the indirect evidential is the more marked term in its semantics. Given Montrul's (2009) finding

of Mood distinctions being more strongly eroded than non-modal inflectional distinctions in Spanish heritage speakers, we expected bilinguals' sensitivity to indirect evidential markers to be more reduced than their sensitivity to direct evidential markers. Difficulty with indirect evidentials is also what the Interface Hypothesis predicts. According to this hypothesis, bilinguals tend to have problems with integrating information from multiple linguistic levels at the syntax-discourse interface and thus should show more difficulty processing marked compared to unmarked forms (e.g., Sorace and Serratrice, 2009). However, both early and late bilinguals were more accurate and quicker to respond to the more marked term (the indirect evidential) here, whose use is licensed only by the availability of a specific type of evidence, than to the less marked term (the direct evidential) in the current study.

Alternatively, we may be able to account for our findings by assuming that, even though Turkish heritage speakers are aware of the semantic and pragmatic properties of indirect evidentials, the direct evidential morpheme -DI has become the default form for referring to past events regardless of information source. That is to say that the bilingual participants take the direct evidential to be a past tense marker without any specific evidential content, whilst they retained the indirect evidential as an evidential form associated with reporting non-witnessed events. This hypothesis broadly fits with Arslan et al. (submitted) finding that early bilingual speakers of Turkish were largely insensitive to mismatches between evidential verb forms and evidential contexts but had retained sensitivity to incorrect tense forms. Although the early bilinguals examined by Arslan et al. (submitted) seemed unable to identify information source violations for either of the two evidential forms, Arslan and Bastiaanse (2014) found an asymmetrical substitution error pattern. The early bilingual speakers of Turkish mistakenly produced direct evidential forms in contexts where an indirect evidential would normally be required. This indicates that the early bilinguals ignored the evidential content of direct evidential forms, using these forms to refer to the past irrespective of whether or not its use was licensed by the type of evidence available. This is also supported by the current findings. When given a visual depiction of directly witnessed evidence for an event, bilingual speakers of Turkish have more problems processing direct evidential forms than monolinguals, whereas they are no different from monolinguals in their processing of indirect evidentials accompanied by a visual depiction of indirect (inferential) evidence.

Recall that one idea behind the conceptual design of this study was to reveal whether and when speakers of an evidential language consider the evidence during processing grammatical evidentiality. That is, we were also interested in whether the speakers were aware of the evidential implications signaled by the verbal forms. Both the behavioral and eye-movements data point in the same direction: both late and early bilinguals fixated less frequently on the target picture in the direct than in the indirect evidential condition, whereas the monolinguals showed no difference between these two conditions in the main time

window. Fewer looks to the target picture in the direct evidential condition means that the bilingual participants fixated more often on the context picture in the direct than in the indirect evidential condition in both the main and late time windows. They also clicked on the context picture more frequently in the direct evidential condition, as shown by their reduced response accuracy. This was not what the monolinguals did. In the late time window, although the monolinguals tended to look at the target picture slightly more often in the direct evidential than the indirect evidential condition, they were equally able to choose the target picture in both conditions. This indicates that the bilinguals were less likely to recognize that the context pictures merely provided a form of evidence, and more likely to mistake the context picture for the target picture, in comparison to the monolinguals.

The time course of participants' eye-movements during processing direct evidentials also differed between the monolingual and bilingual Turkish speakers. The monolinguals shifted their gaze toward the context picture, where the action was shown to be in progress, in the late time window (from about 1200 ms) while processing direct evidentials. This suggests that increased looks toward the context picture allowed the monolinguals to verify that the action could indeed be 'witnessed' directly, compatible with the use of a direct evidential form. This shift was less prominent in the two bilingual groups, although their fixations also changed over time in the late time window due to larger oscillations between the two pictures (see **Figure 3**), indicating that the bilinguals felt less of a need to 'witness' the action, and thus, to verify whether the use of a direct evidential was warranted. This suggests that the direct evidential has been subject to semantic or pragmatic 'bleaching' in Turkish heritage grammars, making it appropriate for use in both 'witnessed' and 'non-witnessed' types of evidential contexts. Examples of a restructuring of grammatical systems in bilingual speakers of minority languages (i.e., heritage speakers) are not in fact uncommon. Polinsky (2006), for instance, reports simplifications in the gender and aspect systems of Russian heritage speakers, and Kim et al. (2009) observed a simplification of the pronominal system in Korean heritage speakers. However, whether or not the apparent erosion of evidentiality distinctions in Turkish heritage speakers is triggered by prolonged exposure to the majority language of our bilingual participants cannot be determined in the absence of a bilingual comparison group whose L2 is typologically different from German (and Dutch).

Conclusion

Our results show that both early and late Turkish/German bilinguals differed from Turkish monolinguals in their processing of direct (but not indirect) evidentiality. These data do not support the Regression Hypothesis or the Interface Hypothesis. We have argued that our findings can be accounted for by assuming that the bilinguals take the direct evidential to be the 'unmarked' default form for referring to past events, in line with what has previously been reported by Arslan and Bastiaanse (2014) and Arslan et al. (submitted). Taken together, our findings from the production, off-line comprehension and online processing of evidentiality by Turkish-German and Turkish-Dutch bilinguals provide converging evidence suggesting that the grammar of evidentiality in these bilinguals has simplified at the representational level. The bilinguals under study are, however, aware that the use of indirect evidential forms is linked to a particular type of evidence, as both our behavioral and eye-movement data suggest that the early and late bilinguals interact with the indirect evidence in a similar way as the monolinguals.

Author Contributions

Conception or design of the experiment: SA, RB, CF; acquisition and analysis of the data: SA; drafting of the manuscript: SA; revising of the manuscript: RB, CF; final approval of the content: SA, RB, CF; agreement to accuracy or integrity of any part of the work: SA, RB, CF.

Acknowledgments

References

Aikhenvald, A. Y. (2004). *Evidentiality*. Oxford: Oxford University Press.

Aksu-Koç, A. (1988). *The Acquisition of Aspect and Modality: The Case of Past Reference in Turkish*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511554353

Aksu-Koç, A. (2000). "Some aspects of the acquisition of evidential in Turkish," in *Evidentials: Turkic, Iranian and Neighbouring Languages*, eds L. Johanson and B. Utas (Berlin: Walter de Gruyter), 15–28.

Aksu-Koç, A. (2014). "Acquisition of modality in Turkish," in *Proceedings of the 17th International Conference on Turkish Linguistics*, (France: University of Rouen).

Aksu-Koç, A., and Slobin, D. I. (1986). "A psychological account of the development and use of evidentials in Turkish," in *Evidentiality. The Linguistic Coding of Epistemiology*, eds C. Wallace and J. Nichols (New York, NY: Ablex Publishing Corporation), 159–167.

Aksu-Koç, A., Terziyan, T., and Taylan, E. E. (2014). Input offers and child uptakes: acquiring mood and modal morphology in Turkish. *Lang. Interact. Acquis.* 5, 62–81. doi: 10.1075/lia.5.1.03aks

Albirini, A., Benmamoun, E., and Chakrani, B. (2013). Gender and number agreement in the oral production of Arabic Heritage speakers. *Biling. Lang. Cogn.* 16, 1–18. doi: 10.1017/S1366728912000132

Albirini, A., Benmamoun, E., and Saadah, E. (2011). Grammatical features of Egyptian and Palestinian Arabic heritage speakers' oral production. *Stud. Second Lang. Acquis.* 33, 273–303. doi: 10.1017/S0272263110000768

Arslan, S., and Bastiaanse, R. (2014). "Tense/aspect and evidentiality in narrative speech of Turkish-Dutch bilingual speakers," in *Proceedings of the 17th International Conference on Turkish Linguistics*, (France: University of Rouen).

Baayen, R. H. (2008). *Analyzing Linguistic Data*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511801686

Barr, D. J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *J. Mem. Lang.* 59, 457–474. doi: 10.1016/j.jml.2007.09.002

Bastiaanse, R., Jonkers, R., and Thompson, C. K. (2008). *Test for Assessing Reference of Time (TART)*. Groningen: University of Groningen.

Chafe, W., and Nichols, J. (1986). *Evidentiality: The Linguistic Coding of Epistemology*. Norwood, MA: Ablex Publishing.

Cook, V. J. (2003). *Effects of the Second Language on the First*. Clevedon: Multilingual matters.

De Bot, K., and Weltens, B. (1991). "Recapitulation, regression, and language loss," in *First Language Attrition: Structural and Theoretical Perspectives*, eds H. Seliger and R. Vago (Cambridge: CUP), 31–51. doi: 10.1017/CBO9780511620720.003

Gürel, A. (2004). Selectivity in L2-induced L1 attrition: a psycholinguistic account. *J. Neurolinguistics* 17, 53–78. doi: 10.1016/S0911-6044(03)00054-X

Jakobson, R. (1941). *Kindersprache, Aphasie, und allgemeine Lautgesetze*. Uppsala: Almqvist & Wiksell.

Johanson, L. (2006). Indirective sentence types. *Turkic Lang.* 10, 73–89.

Keijzer, M. (2010). The regression hypothesis as a framework for first language attrition. *Biling. Lang. Cogn.* 13, 9–18. doi: 10.1017/S136672890999 0356

Kim, J.-H., Montrul, S., and Yoon, J. (2009). Binding interpretations of anaphors by Korean heritage speakers. *Lang. Acquis.* 16, 3–35. doi: 10.1080/10489220802575293

Köpke, B. (2004). Neurolinguistic aspects of attrition. *J. Neurolinguistics* 17, 3–30. doi: 10.1016/S0911-6044(03)00051-4

Köpke, B., and Schmid, M. S. (2004). "Language attrition: the next phase," in *First Language Attrition: Interdisciplinary Perspectives on Methodological Issues*, eds M. Schmid, M. Keijzer, and L. Weilemar (Amsterdam: John Benjamins Publishing Company), 1–46. doi: 10.1075/sibil.28.02kop

Köpke, B., Schmid, M., Keijzer, M., and Dostert, S. (2007). *Language Attrition: Theoretical Perspectives*. Amsterdam: John Benjamins Publishing. doi: 10.1075/sibil.33

Lazard, G. (2001). On the grammaticalization of evidentiality. *J. Pragmat.* 33, 359–367. doi: 10.1016/S0378-2166(00)00008-4

Montrul, S. A. (2002). Incomplete acquisition and attrition of Spanish tense/aspect distinctions in adults bilinguals. *Biling. Lang. Cogn.* 5, 39–68. doi: 10.1017/S1366728902000135

Montrul, S. A. (2008). *Incomplete Acquisition in Bilingualism: Re-Examining the Age Factor*. Amsterdam: John Benjamins Publishing. doi: 10.1075/sibil.39

Montrul, S. A. (2009). Knowledge of tense-aspect and mood in Spanish heritage speakers. *Int. J. Biling.* 13, 239–269. doi: 10.1177/1367006909 339816

Öztürk, Ö, and Papafragou, A. (2007). "Children's acquisition of evidentiality," in *Proceedings from the 31st Annual Boston University Conference on Language Development*, (Boston, MA: Boston University).

Öztürk, Ö., and Papafragou, A. (2008). "Acquisition of evidentiality and source monitoring," in *Proceedings from the 32nd Annual Boston University Conference on Language Development* (Somerville, MA: Cascadilla Press).

Pavlenko, A. (2004). "L2 influence and L1 attrition in adult bilingualism," in *First Language Attrition: Interdisciplinary Perspectives on Methodological Issues*, eds M. Schmid, B. Köpke, M. Keijzer, and L. Weilemar (Amsterdam: John Benjamins Publishing), 47–59.

Plungian, V. A. (2001). The place of evidentiality within the universal grammatical space. *J. Pragmat.* 33, 349–357. doi: 10.1016/S0378-2166(00)00006-0

Polinsky, M. (2006). Incomplete acquisition: american russian. *J. Slav. Linguist.* 14, 191–262.

Polinsky, M. (2008). "Heritage language narratives," in *Heritage Languages: A New Field Emerging*, eds D. Brinton, O. Kagan, and S. Bauckus (New York, NY: Routledge), 149–164.

Rayner, K., Slowiaczek, M. L., Clifton, C., and Bertera, J. H. (1983). Latency of sequential eye movements: implications for reading. *J. Exp. Psychol. Hum. Percept. Perform.* 9, 912–922. doi: 10.1037/0096-1523.9.6.912

R-Core-Team. (2012). *R: A Language and Environment for Statistical Computing*. Vienna: R Core Team.

Seliger, H. W., and Vago, R. M. (1991). *First Language Attrition*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511620720

Sezer, E. (2001). "Finite inflection in Turkish," in *The Verb in Turkish*, ed. E. Erguvanlı-Taylan (Amsterdam: John Benjamins), 1–46.

Slobin, D. I., and Aksu, A. A. (1982). "Tense, aspect and modality in the use of the Turkish evidential," in *Tense-aspect: Between Semantics and Pragmatics*, ed. P. J. Hopper (Amsterdam: John Benjamins), 185–200.

Sorace, A. (2000). Syntactic optionality in non-native grammars. *Second Lang. Res.* 16, 93–102. doi: 10.1191/026765800670666032

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. *Linguist. Approaches Biling.* 1, 1–33. doi: 10.1075/lab.1.1.01sor

Sorace, A., and Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Lang. Res.* 22, 339–368. doi: 10.1191/0267658306sr271oa

Sorace, A., and Serratrice, L. (2009). Internal and external interfaces in bilingual language development: beyond structural overlap. *Int. J. Biling.* 13, 195–210. doi: 10.1177/1367006909339810

Tsimpli, I., Sorace, A., Heycock, C., and Filiaci, F. (2004). First language attrition and syntactic subjects: a study of Greek and Italian near-native speakers of English. *Int. J. Biling.* 8, 257–277. doi: 10.1177/13670069040080030601

Ünal, E., and Papafragou, A. (2013). "Linguistic and conceptual representations of inference as a knowledge source," in *Proceedings from the 37th Annual Boston University Conference on Language Development* (Boston, MA: Boston University).

Willett, T. (1988). A cross-linguistic survey of the grammaticization of evidentiality. *Stud. Lang.* 12, 51–97. doi: 10.1075/sl.12.1.04wil

Yağmur, K. (1997). *First Language Attrition among Turkish Speakers in Sydney*. Tilburg: Tilburg University Press.

New Structural Patterns in Moribund Grammar: Case Marking in Heritage German

Lisa Yager[1], Nora Hellmold[2], Hyoun-A Joo[2], Michael T. Putnam[2], Eleonora Rossi[3], Catherine Stafford[1] and Joseph Salmons[1]*

[1] Center for the Study of Upper Midwestern Cultures, University of Wisconsin–Madison, Madison, WI, USA, [2] Department of Germanic & Slavic Languages and Literatures, Pennsylvania State University, University Park, PA, USA, [3] Psychology and Sociology Department, California State Polytechnic University, Pomona, CA, USA

Research treats divergences between monolingual and heritage grammars in terms of performance—'L1 attrition,' e.g., lexical retrieval—or competence—'incomplete acquisition', e.g., lack of overt tense markers (e.g., Polinsky, 1995; Sorace, 2004; Montrul, 2008; Schmid, 2010). One classic difference between monolingual and Heritage German is reduction in morphological case in the latter, especially loss of dative marking. Our evidence from several Heritage German varieties suggests that speakers have not merely lost case, but rather developed innovative structures to mark it. More specifically, Heritage German speakers produce dative forms in line with established patterns of Differential Object Marking (Bossong, 1985, 1991; Aissen, 2003), suggesting a reallocated mapping of case. We take this as evidence for innovative reanalysis in heritage grammars (Putnam and Sánchez, 2013). Following Kamp and Reyle (1993) and Wechsler (2011, 2014), the dative adopts a more indexical discourse function, forging a tighter connection between morphosyntax and semantic properties. Moribund grammars deploy linguistic resources in novel ways, a finding which can help move us beyond simple narratives of 'attrition' and 'incomplete acquisition.'

Keywords: bilingualism, heritage language, reanalysis, case marking, case syncretism, differential object marking, German

INTRODUCTION

Most research on the grammar of bilinguals known as 'heritage speakers' is framed in terms of what speakers cannot (or can no longer) do, compared to monolingual speakers of their heritage languages, and research typically accounts for these deficiencies in terms of 'incomplete acquisition' and/or 'attrition' (e.g., Benmamoun et al., 2013 and responses to them in the same journal issue). For instance, Montrul et al. (2015, p. 567) summarize research to date as showing that (emphasis added):

> Inflectional morphology, semantics, and the syntax–discourse interface are quite vulnerable to simplification and loss. Several studies of different heritage languages that used different methodologies have shown that HERITAGE SPEAKERS DO NOT MASTER CASE …

Here, we seek to reorient discussions away from that focus on lack or loss and toward understanding heritage grammars in terms of active reanalysis, in line with some other work

on early bilinguals (e.g., Kupisch and Barton, 2013) as well as similar arguments made for non-sequential bilinguals across the lifespan (Putnam and Sánchez, 2013). We reinterpret a classic example of 'loss' in a heritage grammar as an innovative reanalysis on the part of heritage speakers. That example is 'case' in diasporic varieties of German. Many German varieties have three nominal cases (nominative, accusative, dative) and one common scenario is that morphological dative, historically present across Germanic and still present in Standard German (SG) and other varieties, appears to be lost, leaving a nominative-oblique system. This shift has happened in European varieties and heritage varieties. It is exemplified in (1), from Wisconsin Heritage German:

(1) Wisconsin Heritage German (WHG) case marking
(a) Standard-like dative

	WHG	Standard German
	<u>im</u> Boom	im Baum
	in-the-DAT tree	in-the-DAT tree

(b) Accusative for SG dative

	von <u>ein</u> Dorf	von einem Dorf
	from a-NOM-ACC village	from a-DAT village

(c) Innovative marking

	es war in <u>den</u> Haus	im Haus
	it was in the-ACC house	in the-DAT house

Example (1a) reflects that dative is not entirely lost in these varieties, while (1b) exemplifies a morphologically ambiguous form, presumably an accusative in this context, though surface-identical with the nominative form for neuters. As shown in (1c), we also find some innovative marking, in this case a form, *den*, that would be distinctly accusative for a masculine but used here with a neuter noun, which would show no distinction in the standard, as just noted.

Patterns of case reduction have also been observed in other heritage languages (e.g., Russian in Polinsky, 1995, Hindi in Montrul et al., 2012, and comparatively across Spanish, Hindi, and Romanian in Montrul et al., 2015). We present data from three different contact settings and five German varieties in total that show dative marking that differs from canonical three-case systems.

Previous analyses have treated such changes both in terms of failure to acquire case morphology and/or loss through attrition. 'Incomplete acquisition' (Montrul, 2008), understood essentially as the arrested development of certain features of the heritage language (see below), is an unlikely culprit in this process since most speakers in the present study were monolingual speakers of German until around age six, well after when dative would have normally been learned, around age three (Eisenbeiss et al., 2009). Attrition, taken as the loss of some structural property after it has been successfully acquired, would then seem like the obvious source of case loss.

However, closer analysis suggests a more nuanced view, namely, that speakers are developing patterns of Differential Object Marking (DOM), following a hierarchy in which preferences are shown cross-linguistically for marking case on animate and definite arguments over inanimate and indefinite

ones. Aissen (2003, p. 435) defines it this way: "It is common for languages with overt case-marking of direct objects to mark some objects, but not others, depending on semantic and pragmatic features of the object." In the literature, DOM effects are often expressly restricted to DIRECT objects, though the literature since Bossong (1991) has treated complex interactions involving dative objects. As Aissen (2003, p. 446) writes, "In a number of the languages ..., accusative case in a DOM system is identical to dative case ..." In Spanish, for instance, the DOM marker, 'personal *a*,' is also used for indirect objects, and in Hindi *-ko* marks DOM on direct objects but also indirect objects (Montrul et al., 2015, p. 570). Here, dative case marking is retained more often on pronouns than on determiners and, in some varieties, more on definites than indefinites. On the empirical side, this is the first time to our knowledge that the EMERGENCE of new DOM effects has been described for heritage languages. More detailed discussion of dative DOM is left for future work.

Changes in morphological case marking, based on these results, should not simply be viewed as a loss of inflectional morphology but rather need to include the emergence of new semantic-morphosyntactic mapping strategies. Our general conclusion is that heritage bilingual grammars are complete grammatical systems that show structural innovations of the sort we expect in any living language. The patterns we observe are understandable in terms of reanalysis of structural systems (e.g., Polinsky, 2011; Putnam and Sánchez, 2013), and this discussion begins to move research toward modeling the actual implementation.

The question of whether particular 'vulnerable domains' exist in developing bilingual grammars has been pursued in previous studies (e.g., Paradis and Genesee, 1996; Hulk and Müller, 2000; Meisel, 2001; Müller and Hulk, 2001). A primary focus of this research has been on whether or not some aspects of morphosyntax may be affected by interdependent developments rather than the entire grammar system. The general consensus argues for interdependence primarily except for when the grammar interacts with other cognitive (i.e., extra-grammatical) interfaces.

The rest of this paper is structured as follows. The next section gives a brief overview of German case and apparent case reductions in heritage German ('speech islands') and for Germanic more generally. §2 introduces 'incomplete acquisition' and 'attrition' as they have been applied to reductions in inflectional morphology among heritage language speakers, along with data on L1 German case acquisition. §3 presents methods and data from a set of heritage German varieties: §3.1 for Texas German, §3.2 for three varieties from Wisconsin and §3.3 for some initial data on Misionero German (MG) from South America. §4 concludes.

CASE MARKING AND CASE REDUCTION IN GERMANIC AND HERITAGE GERMAN

While SG has a four-case system, the genitive is not widely used in colloquial varieties either historically or today; moreover, genitive case was likely present in heritage varieties only through

exposure in school or reading formal texts for most, so that discussion of Heritage German case best starts from a three-case system, consisting of nominative, accusative, and dative.[1] Case is marked on many pronominal forms and on determiners, though there is considerable syncretism in some paradigms. **Table 1** shows examples of three pronominal and three definite article paradigms drawing on two of German's three genders, masculine and feminine.

The distinction between structural and lexical case in German is debated and here we follow Eisenbeiss et al. (2009, pp. 9–10), who treat accusatives (as either direct objects or complements of prepositions) and datives in the function of indirect object as structural. Dative forms appearing as complements of prepositions or with verbs that govern the dative (*helfen* 'to help', *antworten* 'to answer') are considered lexical. As reviewed by Eisenbeiss et al. (2009), alternatives and variants include views that treat all datives as lexical (Haider, 1985; Haegeman, 1991), that treat all prepositional case use as lexical (Haegeman, 1991; Heinz and Matiasek, 1994), and that treat prepositional datives as structural and accusatives as lexical (Bierwisch, 1988).

(2) **Structural vs. lexical case, after** Eisenbeiss et al. (2009), **focusing on datives**
    Structural
        Nominative and accusative on direct objects.
            ich glaube 'I believe', **sie** arbeitet 'she works'
            sie sieht **mich** 'she sees me', wir kennen **den** Mann 'we know the man'
        Dative on indirect objects:
            er gibt es **denen** 'he gives it to them', sag **mir** etwas 'tell me something'
    Lexical
        Dative with complements of prepositions:
            mit **mir** 'with me', nach **dem** Film 'after the movie'
        Dative with '2Prep'[2] (locative)
            in **der** Schule sein 'to be in school', auf **dem** Bett liegen 'to lay on the bed'
        Dative with 'dative verbs':
            hilf **mir** 'help me', gehört **ihr** 'belongs to her'

Transitive verbs that govern the dative require an object in dative case. This means that the case of the direct object is item-based and not structural. In contrast, ditransitive verbs require a direct object in accusative case and an indirect object in dative case. A simple transformation task illustrates the difference:

(3) **The syntactic distinctiveness of 'dative verbs'**
(a) Ich   sehe   den   Mann.   Der   Mann   wird   gesehen.
    I-NOM   see   the-ACC man   the-NOM man   is   seen
(b) Ich   helfe dem   Mann.   Dem   Mann   wird   geholfen.
    I-NOM   help   the-DAT man   the-DAT man   is   helped
                                    *Der   Mann   wird   geholfen.
                                    the-NOM man   is   helped

[1]We leave aside here varieties that have only two cases, e.g., most dialects of Low German.

[2]We use '2Prep' to refer to prepositions that govern dative or accusative, the former for locative and the latter for motion across boundaries.

TABLE 1 | Example nominal paradigms for German case.

	Pronouns			Determiners		
	1 sg.	2 sg.	3 fem. sg.	masc. sg.	fem. sg.	Plural
Nominativ	ich	du	sie	der	die	die
Accusative	mich	dich	sie	den	die	die
Dative	mir	dir	ihr	dem	der	den(en)

In the case of a verb that governs the dative the direct object cannot be promoted to the subject position in a passive sentence.

A cross-linguistically common pattern of case marking is DOM. Following Aissen (2003), DOM occurs in languages with overt case marking where some direct objects are marked and others are not. What governs DOM is dependent on semantic and pragmatic contexts. Though DOM has not been widely discussed for Germanic, the phenomenon has been the focus of numerous functional, formal, and hybrid perspectives (Lazard, 1984; Bossong, 1985, 1991; de Hoop, 1992; Aissen, 2003; Naess, 2004; de Swart, 2007; Dalrymple and Nikolaeva, 2011). Dalrymple and Nikolaeva (2011, p. 2) argue that "marked objects are associated with the information-structure role of **topic**. The association may be either synchronic or historical. Where the direct connection between marked objects and topicality has been lost through grammaticalization, marked objects in some languages become associated with **semantic features** typical of topics (animacy, definiteness, specificity)." While many architectural and operational differences exist across contemporary linguistic formalisms, we adopt Dalrymple and Nikolaeva's position.

It has often been observed that pronouns retain dative markings longer than determiners or noun phrases, e.g., in the history of English (Lass, 1992, p. 140ff.), but the same pattern is found across various languages undergoing case loss, including Romance, where Spanish, French, and Italian no longer show case in noun phrases but typically retain nominative-oblique and often other forms in pronouns (e.g., Spanish first singular *yo, me, mío(s)/mía(s), mí, conmigo*). For diasporic varieties of German, Rosenberg (2005, p. 230) describes things this way:

> German-speaking language islands also share another striking feature which may result from an internal typological drift common to all German varieties or even to all Germanic and other Indo-European languages: while case reduction in the nominal paradigms is extensive, it is not in the pronominal paradigms. Personal pronouns frequently have a three-case system or retain at least the dative, which includes the possibility of marking the direct-indirect object relation (by common case vs. dative).

This retention of dative marking on pronouns over determiners has been accepted as a pattern, but not placed in a broader context. DOM effects, we propose, play a very different role in Heritage German: Ostensible loss of dative can be better seen as reanalysis of old morphological/syntactic case marking into a new system of variable DOM. DOM has, in fact, been described as "syntactic rules conditioned by semantic factors" (Baerman, 2008, p. 229). None of the long tradition of diachronic research on Germanic case reductions just mentioned

discusses DOM and case loss at all to our knowledge. If there are DOM effects in Heritage German realizations of the SG dative, this leads to some easily testable predictions:

- Pronouns should show case marking over full NPs, e.g., *mit mir* but *mit **den** Mann*;
- Definite should show case marking over indefinite, e.g., less dative on *ein*-determiners (indefinite) than *der*-determiners (definite), so standard *einem* should be realized as *ein/einen* more often than *dem* as *der/den* or *das*.
- Animate should show case marking over inanimate, so humans and animals should show more dative determiners than physical objects.

The ongoing historical loss of morphological case in Germanic languages is reconstructible since the transition from Indo-European to Proto-Germanic. It has been intensely studied for decades from almost every conceivable perspective (see Bousquette and Salmons, forthcoming, or specifically on German, Salmons, 2012). Diasporic German dialects, 'language islands,' show especially widespread patterns of case change, especially dative. This is reported for varieties spoken in Eastern Europe, Brazil, Australia, South Africa, and across North America (see, among many others, Rosenberg, 1994, 2005; Nützel and Salmons, 2011).

Barðdal and Kulikov (2008, p. 470) review various scenarios for case reduction, including phonetic-phonological, morphological and syntactic-semantic accounts, noting that case loss is "typically preceded by a period of variation and alternation between case forms or argument structures." Language contact clearly correlates with loss of inflectional morphology (O'Neil, 1978; Maitz and Németh, 2014). This is one of the most robust findings across myriad dialects and contact settings for heritage German varieties. And as already noted, the pattern extends far beyond Germanic. Benmamoun et al. (2013, p. 142) state: "Morphological deficits in heritage languages are asymmetric; they seem to be more pronounced and pervasive in nominal morphology than in verbal morphology." We turn now to the two major accounts of this pattern.

EXPLAINING REDUCTION OF INFLECTIONAL MORPHOLOGY: INCOMPLETE ACQUISITION AND ATTRITION

As previously noted, the two main accounts of morphological reduction in heritage grammars involve incomplete acquisition and attrition. We treat each in turn after a word about the acquisition of case.

The basic picture of how functionally monolingual L1 learners acquire case proceeds as follows, according to Mills' (1985, p. 155) classic study (confirmed by much research since, which we will not review here):

> The marking of case in the nominative and accusative is only apparent in the masculine gender paradigm. The distinctive marking of nominative and accusative is sporadic before age 3;0;

otherwise the nominative case form is used. This can probably be attributed to an attempt to regularize the paradigm since in the feminine, neuter, and plural paradigms there is no distinction. Dative case appears around age 3;0 and is usually marked correctly except after prepositions. Genitive case does not appear marked on the article in any of the data reported … .

Prepositions start to appear regularly, predominantly in locative use, around age 3;0. Accusative case is frequently overgeneralized after prepositions. This is probably due to the easy confusion of *n* (marking accusative) and *m* (marking dative) in the masculine gender paradigm. From experimental evidence the stative meaning appears to be learned before directional meaning with those prepositions which can have both meanings.

Eisenbeiss et al. (2009) compare two groups of children, a set of typically developing (TD) children and a set of children with Specific Language Impairment (SLI), the former aged 2;6-3;6 at the time of recording and the latter 5;8-7;11. For both groups, structural case was highly accurate and lexical datives, either with prepositions or verbs, were about half dative and half accusative. They also note that case marking is often omitted on what they call '*ein*-determiners': indefinite articles, possessive pronouns, and the negation element *kein-* 'no'. We will pick up on this again below.

Turning now to incomplete acquisition, it is a concept that receives much attention but which often remains ill-defined and poorly understood. Montrul (2008, p. 21), whose treatment of this topic is perhaps the most detailed available, understands incomplete acquisition as "(for lack of a better term) … a mature linguistic state, the outcome of language acquisition that is not complete or attrition in childhood. Incomplete L1 acquisition occurs in childhood, when, for different reasons, some specific properties of the language do not have a chance to reach age-appropriate levels of proficiency after intense exposure to the L2 begins." According to this definition, language acquisition is truncated—incomplete—in bilingual speakers whose developing L1 grammar receives insufficient input (from the standpoint of quantity and/or quality of input) during the formative earlier years of language acquisition (i.e., prior to puberty for Montrul, but see Paradis, 2009 on dating it much earlier, to 2–5 years). The concern is reinforced by Meisel et al. (2013, p. 149) that "the notion of 'incomplete acquisition' is not defined with the desirable precision in the literature on heritage languages." Other views exist, such as those of Pascual y Cabo and Rothman (2012) and Putnam and Sánchez (2013), that heritage grammars are completely acquired grammars, yet distinct from those of other monolingual and bilingual speakers.

It is very unlikely that the emergence of DOM effects in the varieties of diasporic Heritage German we investigate here stems from insufficient input during L1 acquisition or an inability of the speakers to convert this input into intake such that it is integrated into the developing grammar.

Bentz and Winter (2013, p. 18) argue that languages with more L2 speakers, i.e., languages that are used by many speakers who have learned them as a second language, show more case loss than languages with fewer L2 speakers. This fits with evidence that L2 acquisition of case is difficult even under the

best of circumstances. They extend their discussion to language 'enclaves,' using an example from a variety of Heritage German, the one which will provide our first case study below:

> a common finding is that inflectional paradigms are maintained in the first generations after immigration, but in the following generations morphological systems are quickly simplified ... For example, in Texas German, use of the dative went down from 64 to 28.5% (Salmons, 1994, p. 61) within only one generation. This dramatic change happened when ... a considerable number of parents (Boas, 2009, p. 349) decided not to speak Texas German with their children. Thus, the children of this variety successively became L1 speakers of English and L2 learners of Texas German ... This opens up the possibility that case loss is at least partly due to imperfect L2 learning.

Boas does not actually claim that the last generation of Texas German speakers was L2 learners. L2 learning of heritage varieties is rare, and this view seems to reflect basic misunderstandings about heritage languages (cf. Rothman and Treffers-Daller, 2014). Salmons (1994) actually associates the decline in dative marking with the loss of exposure to SG when schools switched from German- to English-medium instruction.

In addition to incomplete acquisition, much literature centers on L1 attrition, referring to a decline in performance-based (vs. competence-based) attributes of a grammar that have been completely acquired. As Montrul (2008, p. 65) clarifies, "attrition in adults affects primarily performance (retrieval, processing, and speed), but does not result in incomplete or divergent grammatical representations." This definition is more or less consistent with other definitions of attrition, such as the one provided by the Oxford English Dictionary (OED Online): "the gradual disappearance of a linguistic feature from a language. Later also: the gradual decline in use of or loss of ability in a language, esp. in a bilingual or multilingual community." With that background, we now turn to data from three varieties of Heritage German which can then be considered in the terms of this discussion.

DATA FROM HERITAGE GERMAN

This section illustrates variable realizations of accusative and especially dative forms across several varieties and regions: in Texas German, in three varieties spoken in eastern Wisconsin, and in a variety of German spoken in South America. 'Dative loss' has often been treated in the black-and-white terms that the name suggests. The first dataset is a reanalysis of old data, while the second comes from work in progress and first reported here and the third set is a first exploration undertaken specifically for this project. Note that these are not the typical heritage speakers discussed in recent research, but instead bilinguals whose families have been, as described below, speaking German varieties in societies with other dominant languages for several generations.

We begin to add some nuance here, first in relatively familiar ways, like realization of dative on pronouns vs. full noun phrases, but then extending to definiteness and animacy.

Texas German

As described in many works, most extensively in Gilbert (1972), German speakers settled in especially central Texas. The settlement was chronologically relatively compact, starting in the 1840s and the language was transmitted over generations until the late 20th century.

Salmons (1994) provides an analysis of Texas German data, based on Gilbert's (1972) *Atlas*, where a set of sentence translations involved what would be SG dative forms, e.g., 'he came with me', cf. SG *mit mir* (dative) and 'he's already in the room,' Standard *im Zimmer*. The first point was to establish that the dative had in fact once been widespread in Texas. **Table 2** below presents that data, showing a rapid and sharp decline in the use of dative among Gilbert's consultants born after about 1911, which Salmons attributes to the removal of German as the medium of instruction in schools around that time.

Further analyses in Salmons (1994) were focused on speakers from particular regions. **Table 3** presents the numbers there, rearranged for our purposes to capture Eisenbeiss et al.'s (2009) distinction between lexical and structural case, discussed above. While Eisenbeiss et al. (2009) found that L1 acquirers mastered structural case quickly and lexical case only later, Texas German adults do not show parallel patterns: The lowest rates of dative are found with prepositions that can either govern dative or accusative, depending on whether they involve location (dative) or motion across a boundary (accusative).

The clearest correlate of where dative is or is not marked is in fact what element it is marked on. As shown in **Table 4**, use with determiners was strikingly low compared to use with pronouns.

The distinction between lexical and structural case, the observations from which this study ultimately grows, is suggestive of the DOM patterns discussed above, where pronouns are at the top of the DOM hierarchy. Let us turn to data from Wisconsin.

Three Wisconsin Communities

A large number of German-speaking settlers arrived in Wisconsin in the latter two-thirds of the 19th century. Unlike the settlement patterns for Texas German, in Wisconsin immigrants from similar geographic, cultural, and linguistic backgrounds often settled together in communities due to social contacts and shared backgrounds, which prevented some contact and supported relatively closed social networks (Frey, 2013, pp. 119–120, and elsewhere). SG also played a role in these communities; members were often fluent in both a dialect and a kind of High German, mutually intelligible with the standard language. The speech of Wisconsin Heritage German (WHG) speakers

TABLE 2 | Texas German dative vs. accusative for standard dative, regional/age stratification, from Salmons (1994).

Date of birth	NW	WC	SW	NE	Total	Percentage
–1899	20–13	43–29	52–16	29–16	144–74	66.1
1900–1911	21–17	22–15	21–11	17–23	81–66	55.1
1912–	16–60	4–30	20–21	9–12	49–123	28.5

(Regions: NW = Northwest, WC = West Central, SW = Southwest, NE = Northeast).

TABLE 3 | Dative by context (raw numbers).

	Struct.	Lexical, prep		Lexical, 2prep			Lexical, verbal	
	gib ihr 'give her'	mit mir 'with me'	mit ihr 'with her'	über dem Bett 'above the bed'	unter dem Baum 'underneath the tree'	im Zimmer 'in the room'	hilft mir 'helps me'	gehört ihnen 'belongs to hem'
Dative	23	22	15	3	5	6	12	34
Accusative	38	36	23	41	30	22	50	24

TABLE 4 | Case use with ...

	Pro		Det	
Dative	104	35.4%	14	8.3%
Accusative	175	59.5%	150	88.8%

today can be described as a standard-like koiné with dialect features.

Yager (forthcoming) compares case marking on nominal and pronominal tokens by 21 WHG speakers from three distinct communities in eastern Wisconsin. Noun phrases and personal pronouns from semi-structured interviews[3] were categorized and coded based on set characteristics, e.g., gender, number, case, article type, animacy, etc. A total of 5,191 nominal and pronominal tokens were analyzed.

The consultants all learned a German koiné at home as their L1, as described above, and acquired English, typically when they began school. They come from three adjacent but distinct regions in eastern Wisconsin (with seven speakers from each region), which are represented by communities with common social networks and settlement histories. The region known as the Holyland was settled by Catholic immigrants from the Eifel region in western central Germany. Lutherans from Rheinhessen settled in the city of Sheboygan and the surrounding area, while the region around the town of Kiel was settled by Low German speakers. Each of these German dialectal regions is known to deal with the German case marking system in different ways, ranging from a three-case system in Rheinhessen, to a nominative-oblique two-case system in the Eifel region, to a single-case system for nouns in Low German dialects.

Although the settlement histories and baseline dialects vary across the three communities, each group appears to mark case in similar ways, illustrated already in (1) at the outset of this article. **Figure 1** shows the proportion of case-marking on definite NPs by region.[4] Each group produces a similar proportion of SG-like case marking versus non-SG-like case marking, i.e., where an object determiner shows a case-marked form that would not be expected, e.g., for the accusative feminine article, which would be identical to the nominative article in SG. The differences in

[3]The interviews, which took place between 2011 and 2014, were conducted by a group of researchers including Alyson Sewell and transcribed by Alyson Sewell. The interviews were carried out in accordance with the requirements and with the approval of the Institutional Review Board of the University of Wisconsin–Madison, under the protocol "Germanic languages and dialects in Wisconsin" (2013-1639).

[4]Two consultants from the Holyland were excluded from the table because they did not produce any case-marked NPs.
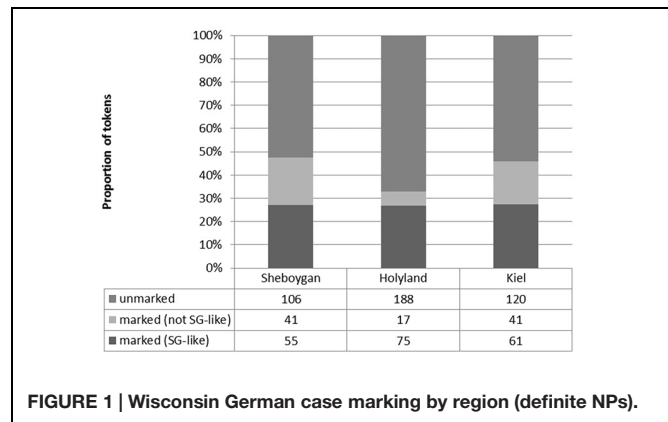

FIGURE 1 | Wisconsin German case marking by region (definite NPs).

case marking between each of these groups are not statistically significant.[5]

With DOM, we would expect to find a higher frequency of case marking on pronouns compared to NPs, as pronouns tend to show a greater degree of both definiteness and animacy. **Figure 2** illustrates these findings for WHG. As **Figure 2** shows, 32.2% of oblique definite NPs are marked in some way, while third person singular pronouns show marking on 41.4% of all tokens. The difference between these two proportions is

[5]Unless otherwise noted, tests of statistical significance are calculated with a two-tailed Z test for two population proportions where $p = 0.05$.
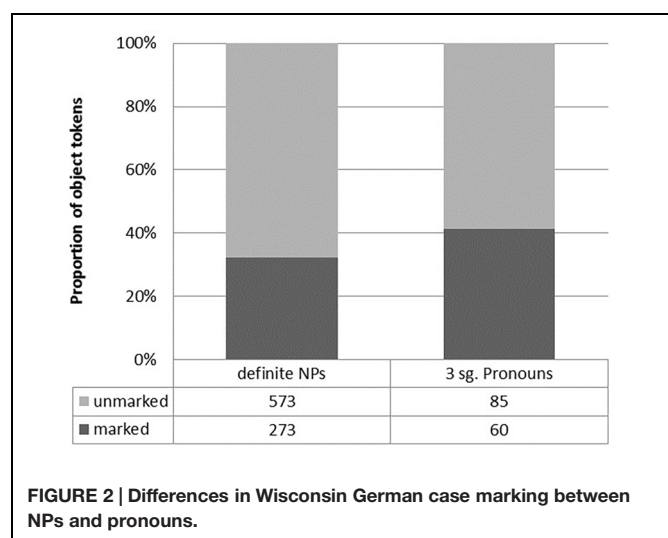

FIGURE 2 | Differences in Wisconsin German case marking between NPs and pronouns.

statistically significant. The overall higher degree of case marking on pronominal tokens is in line with DOM.

There was no significant difference between case marking on animate versus inanimate NPs. However, definite NPs did show a higher frequency of case marking than indefinite NPs. **Table 5** compares case marking on masculine definite and indefinite tokens.

Although the numbers of indefinite tokens are low, the lack of marked indefinite forms compared to the proportion of marked definite forms suggests a correlation between definiteness and case marking in line with DOM.

Not only is the SG case-marking system retained to some extent in each of the three WHG communities, there also appears to be a restructuring of the system around semantic principles, reflecting the emergence of DOM effects.

Misionero German

Misionero German comprises regional dialects of German from the Volga German area spoken in the Misiones province in northeastern Argentina. MG speakers acquired the German variety as their first language (L1). Over time, they have become dominant in their L2 Brazilian Portuguese, the current language of the community, and MG has become moribund. Later, these MG speakers, especially those under the age of 40, acquired Spanish as an L3, which is also widely spoken throughout the Misiones Province. Today, the majority of these transitional trilingual German-Portuguese–Spanish speakers are settled along the upper part of the Uruguay River, from El Soberbio to Panambí. The following data come from speakers in this region (see Putnam and Lipski, forthcoming for an overview).[6]

Free speech data from seven speakers were transcribed and analyzed following the conventions used in Yager (forthcoming), yielding a total of 1,565 tokens; 842 in NP; 283 of these in PP; and 697 pronouns. Because the raw numbers for this first sample are extremely low, not allowing even for use of non-parametric statistics, we report results as descriptive statistics, which will allow describing a general trend in the pattern of performance. The addition of more data in the next phases of this research will be important to confirm the observed trends. First, we looked at differences in case marking between full NPs and pronouns in order to analyze the data for possible DOM effects, as reported in **Figure 3**. Even though the overall number of third person singular pronouns is very small compared to the definite NPs, these preliminary results show that pronouns tend to be marked more frequently (75%) than NPs (53%).

Second, we looked at differences between case marking in definite and indefinite determiners. Even with a small number of tokens, a trend can be seen toward more case marking on definite than indefinite determiners. These results are summarized in **Table 6**. For accusative case marking, only 55% of definite determiners are case-marked. However, 80% of indefinite

TABLE 5 | Wisconsin German case and definiteness.

	Marked	Unmarked	Total tokens
Definite accusative	70	81	151
Indefinite accusative	1	27	28
Definite dative	127	32	159
Indefinite dative	0	3	3
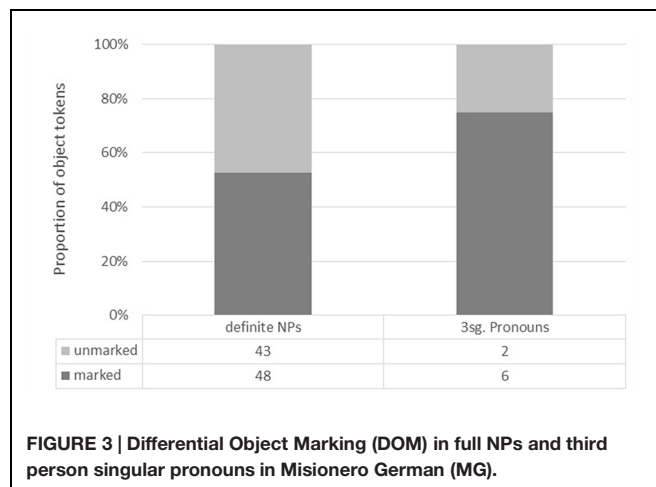
FIGURE 3 | Differential Object Marking (DOM) in full NPs and third person singular pronouns in Misionero German (MG).

TABLE 6 | Trend toward DOM in definite and indefinite determiners in Misionero German.

	Marked	Unmarked	Total
Definite accusative	5 (55%)	4 (45%)	9
Indefinite accusative	1 (20%)	4 (80%)	5
Definite dative	16 (59%)	11 (41%)	27
Indefinite dative	1 (50%)	1 (50%)	2

determiners are unmarked for case. Against this background, a slight trend for DOM of definite determiners can be inferred. Dative case marking shows a similar pattern, with 59% of definite determiners being case-marked.

No DOM effect was found for animate versus inanimate objects. The analysis of the MG data shows DOM with pronouns more marked for case than full NPs, and definite determiners more than indefinite ones. These findings align with the results from WHG.

In summary, one of the most widespread findings in diasporic German has been the loss of case, especially dative marking. Taking a different approach where we examine more nuanced patterns of the realization of dative, a different picture emerges: Across Texas German, three varieties of Wisconsin German and Misionero German, we find distinct but related patterns of case marking, all consistent with dative-based DOM effects.

CONCLUSION

The data presented here point to the emergence of a cross-linguistically familiar generalization in the realization of case marking, namely a particular form of DOM. Traditionally

framed in terms of loss or attrition, these patterns in fact show the development of new grammatical generalizations in these communities. Our findings complicate the traditional narrative of loss and simplification in heritage language grammar, especially with regard to nominal morphology.

The communities analyzed here are geographically very distant from one another, and in contact with different, typologically distinct languages and dialects. In their comparative study of DOM-loss in the English-dominant context of North America, Montrul et al. (2015, p. 566) observe that heritage speakers of Spanish, Hindi, and Romanian "seem to adopt the grammar of English, which does not overtly mark direct objects, and accept non-target sentences with animate, specific direct objects without DOM." The patterns observed in our data, though, cannot be explained simply in terms of direct influence from sociolinguistically dominant L2 grammars, i.e., English, rural vernacular Portuguese, and Spanish. Nor can they reflect spread from one community to another, and because the original input varieties were from different areas and German does not show classic patterns of DOM effects, they are very unlikely to have sprung from seeds imported with initial immigration. Instead, we see a new, divergent grammatical property, the rise of DOM. As is often the case with DOM, its occurrence is tendential rather than categorical.

Appealing to incomplete L1 acquisition as the force behind these changes is not promising, because, as we have noted, German-speaking children develop command of structural case by age 3. We thus should expect children exposed until school age to varieties of German that license dative case to have successfully acquired at least structural datives. All speakers use the dative in a range of grammatical contexts (both structural and lexical), including those with more or less exposure to SG. Similarly, L1 attrition is unlikely since the DOM-patterns we observe are arguably as complex as or more complex than the earlier system. To understand these patterns, we must get past the narratives of "collapse" and "loss" that are commonly attributed to heritage grammars.

In contrast, the patterns we find here are consistent with the position of Putnam and Sánchez (2013), who see heritage grammars as full grammars, capable of change, including reanalysis, in the ways that all grammars are. At the same time, our results also raise issues to be pursued in later work. For instance, how do typological drift and ease-of-processing procedures inform the restructuring process (cf. Hawkins, 2004; Culicover, 2013)? Another challenge regards the connection between more structural units such as morphology and syntax and their relationship to semantics and pragmatics/information structure (see §2.3). Also, our work suggests that variability in heritage grammars should include factors such as age of the speakers, specifically vis-à-vis cognitive functions. Language performance changes with normal aging, as a factor of cognitive changes that occur in normal aging. As Rossi and Diaz (forthcoming) point out, language changes due to normal aging are at times conflated with changes in language processing due to bilingualism and language contact. The populations that were tested in this set of studies exemplify how investigating heritage languages in speakers at different ages (younger adults and older adults) are of importance for future research.

A final question is whether these observed trends occur more broadly across Germanic, past and present. It would be a worthwhile pursuit to explore whether other Germanic languages that have lost case also reorganize their inflectional systems along similar lines.

REFERENCES

Aissen, J. (2003). Differential object marking: iconicity vs. economy. *Nat. Lang. Linguist. Theory* 21, 435–483. doi: 10.1023/A:1024109008573

Baerman, M. (2008). "Case syncretism," in *The Oxford Handbook of Case*, eds A. Malchukov and A. Spencer (Oxford: Oxford University Press), 219–230.

Barðdal, J., and Kulikov, L. (2008). "Case in decline," in *The Oxford Handbook of Case*, eds A. Malchukov and A. Spencer (Oxford: Oxford University Press), 470–478.

Benmamoun, E., Montrul, S., and Polinsky, M. (2013). Heritage languages and their speakers: opportunities and challenges for linguistics. *Theor. Linguist.* 39, 129–181. doi: 10.1515/tl-2013-0009

Bentz, C., and Winter, B. (2013). Languages with more second language learners tend to lose nominal case. *Lang. Dyn. Change* 3, 1–27.

Bierwisch, M. (1988). "On the grammar of local prepositions," in *Syntax, Semantik und Lexikon*, eds M. Bierwisch, W. Motsch, and I. Zimmermann (Berlin: Akademie-Verlag), 1–65.

Boas, H. C. (2009). "Case loss in Texas German: the influence of semantic and pragmatic factors," in *The Role of Semantic, Pragmatic, and Discourse Factors in the Development of Case*, eds J. Barðdal and S. L. Chelliah (Amsterdam: John Benjamins), 347–373.

Bossong, G. (1985). *Empirische Universalienforschung: Differentielle Objektmarkierung in den Neuiranischen Sprachen*. Tübingen: Narr.

Bossong, G. (1991). "Differential object marking in Romance and beyond," *in New Analyses in Romance linguistics*, eds D. Wanner and D. Kibbee (Amsterdam: John Benjamins), 143–167. doi: 10.1075/cilt.69.14bos

Bousquette, J., and Salmons, J. (forthcoming). "The Germanic Languages," in *The Indo-European Languages*, ed. M. Kapović (London: Routledge Publisher).

Culicover, P. (2013). *Grammar and Complexity: Language at the Intersection of Competence and Performance*. Oxford: Oxford University Press.

Dalrymple, M., and Nikolaeva, I. (2011). *Objects and Information Structure*. Oxford: Oxford University Press.

de Hoop, H. (1992). *Case Configuration and Noun Phrase Interpretation*. Ph.D. dissertation, University of Groningen, Groningen.

de Swart, P. (2007). *Cross-Linguistic Variation in Object Marking*. Ph.D. dissertation, Radboud University-Nijmegen, Nijmegen.

Eisenbeiss, S., Bartke, S., and Clahsen, H. (2009). Structural and lexical case in child german: evidence from language-impaired and typically developing children. *Lang. Acquis.* 13, 3–32. doi: 10.1207/s15327817la1301_2

Frey, B. (2013). *Towards a General Theory of Language Shift: A Case Study in Wisconsin German and North Carolina Cherokee*. Ph.D. dissertation, University of Wisconsin – Madison, Madison, WI.

Gilbert, G. G. (1972). *Linguistic atlas of Texas German*. Austin, TX: University of Texas Press.

Haegeman, L. (1991). *Introduction to Government and Binding*. Oxford: Blackwell.

Haider, H. (1985). "The case of german," in *Studies in German Grammar*, ed. J. Toman (Dordrecht: Foris), 65–101.

Hawkins, J. A. (2004). *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199252695.001.0001

Heinz, W., and Matiasek, J. (1994). "Argument structure and case assignment in german," in *German in Head-Driven Phrase Structure Grammar*, eds J. Nerbonne, K. Netter, and C. Pollard (Stanford, CA: Stanford University Press), 199–236.

Hulk, A., and Müller, N. (2000). Bilingual first language acquisition at the interface between syntax and pragmatics. *Bilingualism* 3, 227–244. doi: 10.1017/S1366728900000353

Kamp, H., and Reyle, U. (1993). *From Discourse to Logic*. Dordrecht: Kluwer.

Kupisch, T., and Barton, D. (2013). How bilinguals deal with variation: generic reference in adult German bilinguals. *Stud. Linguist.* 67, 1–27. doi: 10.1111/stul.12006

Lass, R. (1992). "Phonology and morphology," in *The Cambridge History of the English Language*, Vol. 2, ed. N. Blake (Cambridge: Cambridge University Press), 23–155. doi: 10.1017/CHOL9780521264754.003

Lazard, G. (1984). "Actance variations and categories of the object," in *Objects: Towards a Theory of Grammatical Relations*, ed. F. Plank (London: Academic Press), 269–292.

Maitz, P., and Németh, A. (2014). Language contact and morphosyntactic complexity: evidence from German. *J. Ger. Linguist.* 26, 1–29. doi: 10.1017/s1470542713000184

Meisel, J. (2001). "The simultaneous acquisition of two first languages: early differentiation and subsequent development of grammars," in *Trends in Bilingual Acquisition*, eds J. Cenoz and F. Genesee (Oxford: Blackwell), 10–35.

Meisel, J., Elsig, M., and Rinke, E. (2013). *Language Acquisition and Change: A Morphosyntactic Perspective*. Edinburgh: University of Edinburgh Press. doi: 10.3366/edinburgh/9780748642250.001.0001

Mills, A. E. (1985). "The acquisition of German," in *The Crosslinguistic Study of Language Acquisition*, Vol. 1, ed. D. I. Slobin (Hillsdale, NJ: Lawrence Erlbaum), 141–254.

Montrul, S. (2008). *Incomplete Acquisition in Bilingualism: Re-Examining the Age Factor*. Amsterdam: John Benjamins. doi: 10.1075/sibil.39

Montrul, S. A., Bhatt, R. M., and Bhatia, A. (2012). Erosion of case and agreement in Hindi heritage speakers. *Linguist. Approaches Biling.* 2, 141–176. doi: 10.1075/lab.2.2.02mon

Montrul, S., Bhatt, R., and Girju, R. (2015). Differential object marking in Spanish, Hindi, and Romanian as heritage languages. *Language* 91, 564–610. doi: 10.1353/lan.2015.0035

Müller, N., and Hulk, A. (2001). Crosslinguistic influence in bilingual language acquisition: Italian and French as recipient languages. *Bilingualism* 4, 1–21. doi: 10.1017/S1366728901000116

Naess, Å. (2004). What markedness marks: the Markedness Problem with direct objects. *Lingua* 114, 1186–1212. doi: 10.1016/j.lingua.2003.07.005

Nützel, D., and Salmons, J. (2011). Structural stability and change in language contact: evidence from American German. *Lang. Linguist. Compass* 5, 705–717. doi: 10.1111/j.1749-818X.2011.00308.x

O'Neil, W. (1978). The evolution of the Germanic inflectional systems: a study in the causes of language change. *Orbis* 27, 248–286.

Paradis, J., and Genesee, F. (1996). Syntactic acquisition in bilingual children: autonomous or independent? *Stud. Second Lang. Acquis.* 18, 1–15. doi: 10.1017/S0272263100014662

Paradis, M. (2009). *Declarative and Procedural Determinants of Second Languages*. Amsterdam: John Benjamins. doi: 10.1075/sibil.40

Pascual y Cabo, D., and Rothman, J. (2012). The (il)logical problem of heritage speaker bilingualism and incomplete acquisition. *Appl. Linguist.* 33, 450–455. doi: 10.1093/applin/ams037

Polinsky, M. (1995). "American Russian: language loss meets language acquisition," in *Formal Approaches to Slavic Linguistics, Cornell Meeting*, eds W. Brown, E. Dornisch, N. Kondrashova, and D. Zec (Ann Arbor: Michigan Slavic Publications), 370–406.

Polinsky, M. (2011). Reanalysis in adult heritage language. *Stud. Second Lang. Acquis.* 33, 305–328. doi: 10.1017/S027226311000077X

Putnam, M., and Lipski, J. (forthcoming). Null arguments in transitional trilingual grammars: field observations from Misionero German. *Multilingua*. doi: 10.1515/multi-2014-0111

Putnam, M., and Sánchez, L. (2013). What's so incomplete about incomplete acquisition? A prolegomenon to modeling heritage grammars. *Linguis. Approaches Biling.* 3, 478–508. doi: 10.1075/lab.3.4.04put

Rosenberg, P. (1994). "Varietätenkontakt und Varietätenausgleich bei den Russlanddeutschen: orientierungen für eine moderne Sprachinselforschung," in *Sprachinselforschung: Eine Gedenkschrift für Hugo Jedig*, eds N. Berend and K. J. Mattheier (Frankfurt: Peter Lang), 123–164.

Rosenberg, P. (2005). "Dialect convergence in the German language islands (Sprachinseln)," in *Dialect Change: Convergence and Divergence in European Languages*, eds P. Auer, F. Hinskens, and P. Kerswill (Cambridge: Cambridge University Press), 221–235.

Rossi, E., and Diaz, M. (forthcoming). The aging linguistic brain: Bilingualism as an advantage? *Linguistic Approaches to Bilingualism. Special issue*, Bilingualism and Aging.

Rothman, J., and Treffers-Daller, J. (2014). A prolegomenon to the construct of the native speaker: heritage speaker bilinguals are natives too! *Appl. Linguist.* 35, 93–98. doi: 10.1093/applin/amt049

Salmons, J. (1994). "Naturalness and morphological change in Texas German," in *Sprachinselforschung: Eine Gedenkschrift für Hugo Jedig*, eds N. Berend and K. J. Mattheier (Frankfurt: Peter Lang), 59–72.

Salmons, J. (2012). *A History of German: What the Past Reveals about Today's Language*. Oxford: Oxford University Press.

Schmid, M. (2010). *Language Attrition*. Cambridge: Cambridge University Press.

Sorace, A. (2004). Native language attrition and developmental instability at the syntax-discourse interface: data, interpretation and methods. *Biling. Lang. Cogn.* 7, 143–145. doi: 10.1017/S1366728904001543

Wechsler, S. (2011). Mixed agreement, the person feature, and the Index/Concord feature. *Nat. Lang. Linguist. Theory* 29, 999–1031. doi: 10.1177/13670069040080030601

Wechsler, S. (2014). "The structure of Swedish pancakes," in *The Core and the Periphery: Data-Driven Perspectives of Syntax Inspired by Ivan A. Sag*, eds P. Hofmeister and E. Norcliffe (Stanford: Canadian Second Language Institute), 71–98. doi: 10.1007/s11049-011-9149-x

Yager, L. M. (forthcoming). *Morphosyntactic Variation and Change in Wisconsin Heritage German*. Madison, WI: University of Wisconsin – Madison.

CrossMark

Portmanteau Constructions, Phrase Structure, and Linearization

Brian Hok-Shing Chan*

Department of English, Faculty of Arts and Humanities, University of Macau, Macau, China

In bilingual code-switching which involves language-pairs with contrasting head-complement orders (i.e., head-initial vs. head-final), a head may be lexicalized from both languages with its complement sandwiched in the middle. These so-called "portmanteau" sentences (Nishimura, 1985, 1986; Sankoff et al., 1990, etc.) have been attested for decades, but they had never received a systematic, formal analysis in terms of current syntactic theory before a few recent attempts (Hicks, 2010, 2012). Notwithstanding this lack of attention, these structures are in fact highly relevant to theories of linearization and phrase structure. More specifically, they challenge binary-branching (Kayne, 1994, 2004, 2005) as well as the Antisymmetry hypothesis (ibid.). Not explained by current grammatical models of code-switching, including the Equivalence Constraint (Poplack, 1980), the Matrix Language Frame Model (Myers-Scotton, 1993, 2002, etc.), and the Bilingual Speech Model (Muysken, 2000, 2013), the portmanteau construction indeed looks uncommon or abnormal, defying any systematic account. However, the recurrence of these structures in various datasets and constraints on them do call for an explanation. This paper suggests an account which lies with syntax and also with the psycholinguistics of bilingualism. Assuming that linearization is a process at the Sensori-Motor (SM) interface (Chomsky, 2005, 2013), this paper sees that word order is not fixed in a syntactic tree but it is set in the production process, and much information of word order rests in the processor, for instance, outputting a head before its complement (i.e., head-initial word order) or the reverse (i.e., head-final word order). As for the portmanteau construction, it is the output of bilingual speakers co-activating two sets of head-complement orders which summon the phonetic forms of the same word in both languages. Under this proposal, the underlying structure of a portmanteau construction is as simple as an XP in which a head X merges with its complement YP and projects an XP (i.e., X YP → [$_{XP}$ X YP]).

Keywords: code-switching, portmanteau construction, word order, phrase structure, linearization

INTRODUCTION: THE PORTMANTEAU CONSTRUCTION IN BILINGUAL CODE-SWITCHING

This paper seeks a new account of a specific construction in bilingual code-switching which has so far received few in-depth treatments and remains not well-understood, based on existing data gleaned from all works that are accessible, including published papers and unpublished dissertations. The portmanteau construction in code-switching, which involves the juxtaposition

of two synonymous morphemes from two different languages[1], has been attested in various datasets for decades (Nishimura, 1985, 1986; Park, 1990; Sankoff et al., 1990), but nonetheless there had been no systematic studies of the construction (that I know of) until quite recently (Hicks, 2010, 2012). The form of the portmanteau construction is sketched below.

(1) $[_{XP} X_A [YP (ZP)] X_B]$
(Language A is head-initial in XP whereas language B is head-final in XP)

X, the doubled element, is a head, whereas the shared element is the complement of this head, namely, YP. In some cases, a head, such as a ditransitive verb, may select two complements, hence YP and ZP[2]. The languages which participate in the portmanteau construction (i.e., A and B) are mostly typologically different with one (say, language A) being a VO language and another (say, language B) being an OV language, for instance, Japanese-English (Nishimura, 1985, 1986; Azuma, 1993, 1997, 2001; Takagi, 2007; Furukawa, 2008; Namba, 2012a,b), Korean-English (Park, 1990), Hindi-English (Pandit, 1986), Tamil-English (Sankoff et al., 1990), or Marathi-English (Hicks, 2010, 2012), etc. Some data of the portmanteau construction are also attested in a pair of an OV language and a partially VO language (e.g., Dutch-Turkish

in which Dutch is VO in main clause but OV in subordinate clause—see Backus, 1996, 2003), or two SVO languages (e.g., Cantonese-English in which Cantonese is postpositional but English is prepositional—see Chan, 1998, 2015).

It appears that in such constructions the doubled element is often the verb. For Nishimura (1995, p. 167), "'[p]ortmanteau sentences' involve a specific type of repetition: an English sentence and its Japanese equivalent are combined with a commonly-shared constituent. Portmanteau sentences come out in SVOV order: O is the common constituent. The first V is English, and the final V is Japanese." The following are some examples in which the doubled element is a verb whereas the shared element is an object DP.

(2) We **bought** [about two pounds *gurai*] ***kaettekita*** *no*
about   bought   PRT
"We bought about two pounds."
(English-*Japanese*, Nishimura, 1986, p. 139)

(3) One day my friend **brought** [two watch] *kaciyo*
have
wasseyo
come (=bring)
"One day my friend brought two watches."
(English-*Korean*, Park, 1990, p. 103)

There are some data which look like examples (2)–(3), such as (4) and (5) below, but the two verbs actually carry quite different meanings. Strictly speaking, they do not involve the same word being lexicalized in two languages, and therefore they are considered very different and excluded from the present account.

(4) I still have [*etten namca*] *coha-hayyo*
certain man   like
"I still have a certain man that I like."
(English-*Korean*, Park, 1990, p. 103)

(5) You pull [this much] *tsukau desho*
use   will
"You pull this much that you'll use."
(English-*Japanese*, Nishimura, 1986, p. 139)

The shared element of a doubled verb may be more complex than just an object DP. In (6) there are two objects since the doubled verb is ditransitive "*give.*" In (7) the shared element is a PP.

(6) They **gave** [me] [a research grant] ***koDutaa***
give (3-Sg-Past)
"They gave me a research grant."
(English-*Tamil*, Sankoff et al., 1990, p. 93)

(7) I was **talking** [to *oru orutanooDa*] ***peesinDu*** *irunten*
one person   talk-PROG be(1SG-PAST)
"I was talking to one person."
(English-*Tamil*, Sankoff et al., 1990, p. 93)

Some verbs take clauses as complements, and it is not surprising to find the following examples in which the shared element is a CP.

[1]There is an issue of whether there are cross-linguistic synonyms or "translation equivalents" which are really "synonymous" with exactly the same meaning in terms of reference, style or connotations. In the code-switching literature, it has indeed been suggested that code-switching of a single word or a short phrase is motivated by the fact that this word or phrase conveys nuanced meanings which are not conveyed by its synonym or translation (e.g., Li, 2001; Curcó, 2005). In other words, the code-switched word or phrase is incurred by virtue of being the (more) appropriate expression or the "mot juste" (Gafaranga, 2000). Nonetheless, if we consider that the meaning of a word is fuzzy and fluid, being adjusted and fine-tuned in different contexts (Wilson and Carston, 2007), it is plausible that the cross-linguistic synonyms are intended or treated as "exact equivalents" in some contexts. For instance, on fairly formal occasions where only one language is expected, a bilingual speaker may code-switch to a word which is the mot juste (Gafaranga, 2000) but immediately afterwards provides a synonymous translation of that word in the language that is supposed to be the language in the ongoing interaction. In these acts of "medium repair" (Gafaranga, 2000), the code-switched word and its translation are arguably intended and treated as equivalent and identical in meaning, even though in other contexts the bilingual speaker may use the two words to express nuanced meanings. In the context of portmanteau constructions, it is plausible that the doubled words are treated as equivalents with identical meaning. That is, the use of two (near-) synonymous forms from both languages is not so much due to the expression of nuanced meanings, but it is motivated syntactically to satisfy both the head-initial and head-final orders which are co-activated (see below for more details). The synonymy or equivalence of the doubled words tallies with the intuition of many authors of the papers who are also bilinguals interacting with those producing the code-switched, portmanteau constructions, as shown in their glosses and English translations of the examples (e.g., Nishimura, 1985, 1986, 1995; Park, 1990; Namba, 2012a,b). In some rare cases where the doubled words are quite different in meaning and hence not synonymous, these authors were able to notice and indicated that accordingly in the glosses and translations - see (4) and (5). Besides, there is some psycholinguistic evidence that forms of cross-linguistic synonyms or translations are co-activated when a certain meaning (or lemma) is activated (e.g., *dog* and *perro* for a Spanish-English bilingual) to the point that a cross-linguistic synonym (e.g., *perro*) facilitates the access and production of *dog* in picture-word-interference experiments (Costa et al., 2000; Runnqvist et al., 2013).

[2]See example (6) below in which a ditransitive verb is doubled and the shared element consists of two objects or phrases.

(8) Everybody **think** [**that** *nay-ka yenge-lul   cal hanta-**ko**]
            C   I-NOM English-ACC well do-C
    sayngkakhayyo
    think
    "Everybody thinks that I'm a good English speaker."
                                    (English-*Korean*, Park, 1990 p. 103)

(9) Many people **told** [me] [**that** *cey-ka   hankwukcek-ita-**ko**]
                        I-NOM Korean-oriented-C
    malhaysseyo
    told
    "Many people told me that I am Korean-oriented."
                                    (English-*Korean*, Park, 1990, p. 103)

Whereas, the complementizer is also doubled in examples (8) and (9) above, sometimes the complementizer is phonetically realized in one language only, as in (10) below.

(10) I **think** [it's the European influence-*nu*] ***ninakiren***
                        that think (I-SG-PRES)
    "I think that it's the European influence."
                                    (English-*Tamil*, Sankoff et al., 1990, p. 92)

The examples given so far contain lexical verbs. Apart from these verbs, the copula verb "*be*" also takes part in portmanteau sentences, as in examples (11)–(14) below.

(11) Dus in Nederland **zijn**   [zoveel   devlet hastanesi]
    So in Holland   are-3PL so-many state   hospital
            var
            there-are
        "so in Holland there are so many state hospitals…"
                                    (Dutch-*Turkish*, Backus, 1996, p. 348)

(12) It **was** [*cengmal* exiting game] -***iyesseyo***
            really   COP-PAST
    "It was really an exciting game."
                                    (English-*Korean*, Park, 1990, p. 103)

(13) There**'s** [children] ***iru yo***
                    V (existential)
    "There are children."
                                    (English-*Japanese*, Nishimura, 1985, p. 140)

(14) She will not come to me because the hindu system **is**

    [*tarah kaa*] **hai**
    that of is
    "She will not come to me because the Hindu system is like that."

                                    (English-*Hindi*, Pandit, 1986, p. 41)

An auxiliary verb can also be doubled, such as (7) above ("*was*" in English and "*irunten*" from Tamil). A similar example is (15) below, where the doubled element is an auxiliary verb cliticized with the negation marker, and the shared element is a verb or a verb phrase.

(15) My parents **didn't** [*helak-haci*] ***anasseyo***
                        allow-do V+NEG
    "My parents didn't allow (me to do it)."
                                    (English-*Korean*, Park, 1990, p. 104)

The doubled element is not necessarily a verb or an auxiliary verb. Examples (8) and (9) above have already shown that a complementizer may be doubled. In (16) below, it is a subordinator, which is similar to a complementizer in taking a TP or IP complement, that is being doubled.

(16) Just **because** [*avaa innoru*   color and race]
    just because they   different color and race
            engindratunaale
            of-because
    "Just because they are of different color and race."
                                    (English-*Tamil*, Sankoff et al., 1990, p. 93)

In examples (17–20), the doubled element is an adposition.

(17) I could run every you know **in** [thirty minutes]
            madhe once a day.
            in
    "I could run every, you know, in thirty minutes once a day."
                                    (English-*Marathi*, Hicks, 2010, p. 45)

(18) Look for the things she buys **for** [Sean] ***ni***
                                for
    "Look for the things she buys for Sean."
                                    (English-*Japanese*, Nishimura, 1986, p. 140)

(19) **According to** [the schedule] ***paDi***
                        according-to
    "According to the schedule… "
                                    (English-*Tamil*, Sankoff et al., 1990, p. 93)

(20) **After** [*ni1 -go3* review] ***zi1-hau6***…
            DEM CL   after
    "After this review…"
                                    (English-*Cantonese*, Chan, 1998, p. 204)

Referring back to the Japanese-English example (2), there is an English preposition "*about*" that is coupled with the Japanese one "*gurai*" in the object "*about two pounds gurai.*" The prepositions here, however, do not act as prototypical prepositions that mark the location or semantic role of a DP, but rather they somehow modify a noun phrase that denotes an object with a quantity. Accordingly, they look like a "pre-determiner" in traditional descriptive grammar. In generative grammar, they are most probably not instantiations of a P category but more likely of a functional head in the D domain, probably a quantifier head Q. The following is another example, also from Japanese-English, in which a "pre-determiner," probably a Focus head F, is doubled.

(21) Vegas *it-tara dare*   **even** [the tour leader]
            go-if   anyone
            demo they don't lend him money
            even
    "If you go to the Vegas, even the tour leader doesn't lend him money (if somebody has been robbed).
                                    (English-*Japanese*, Furukawa, 2008, p. 286)

Summarizing this survey of portmanteau constructions, we see that heads which take part in the construction include verb, auxiliary verb, preposition (or adposition), complementizer, subordinator, and some functional heads in the DP domain.

Other categories which act as syntactic heads, including noun, adjective, modal verb or conjunction [with a possible exception in (23), see below] have not been found in existing data of portmanteau constructions (see more discussion below).

In all corpora in which the portmanteau construction is found, the bilingual speakers also produce non-portmanteau code-switched constructions; that is, these heads do not have to be doubled [see Chan, 2003, 2008 for a quick survey, also see (39)–(44) below]. In other words, the portmanteau construction is an optional structure.

Repetition involves not only words or free morphemes but also bound morphemes.

(22) …dzimwe dzenguva tinenge tichiita **ma**-*game*-**s** panze
"… sometimes we will be doing games outside."
(Shona-*English*, Myers-Scotton, 1993, p. 132)

Repetition may also take place with a word (or free morpheme) and a bound morpheme which are synonymous. In (23) below, the Spanish conjunction "*pero*" is doubled with "*sti*" from Aymara, which appears to be a bound morpheme affixed to nouns[3] (Hicks, 2010, p. 16).

(23) **pero** *sorro*- **sti**   *wali astuturi*- *tajna*…
but   fox   COORDINATOR very keen   3SG.PRT.EVI
"But the fox was very keen…"
(Spanish-*Aymara*, Stolz, 1996, p. 10,
cited in Hicks, 2010, p. 16)

Covering phenomena illustrated by (22) and (23), Hicks (2010, 2012) counts the portmanteau construction as one type of "morphosyntactic doubling." This paper acknowledges that the process underlying the doubling in (22) and (23) may be very similar to that underlying the doubling of words as shown in examples (2)–(3) and (6)–(21) above. In particular, as a syntactic head combines with its complement in a fixed order (i.e., head-initial or head-final), a bound morpheme is always attached to a root of a particular category (e.g., a plural affix is always attached to a noun, etc.) in a fixed order (i.e., prefix, suffix, or the more uncommon infixes). There are also differences awaiting explanation[4], and existing data of "morphological doubling" (i.e., two synonymous bound morphemes from two languages) are extremely rare, namely, a few instances from Myers-Scotton (1993) as quoted in Hicks (2010, 2012) and marginally (23) above. Hence, this paper focuses on the portmanteau construction or syntactic doubling in code-switching (i.e., two synonymous words or free morphemes from two languages), which does not deny the possibility of pursuing a uniform account of

"morphological doubling" and "syntactic doubling" when more data of the former are uncovered.

The remaining parts of this paper are structured as follows. The next section discusses differences between portmanteau constructions and monolingual syntactic doubling. The following one proposes constraints on the portmanteau constructions. These constructions are then tested against current paradigms of the syntax of code-switching and more general syntactic theories of phrase structure (e.g., Antisymmetry). A new account based on syntax and processing will then be forwarded, followed by a discussion of some residual issues and the conclusions.

PORTMANTEAU CONSTRUCTIONS AND MONOLINGUAL DOUBLING PHENOMENA

Putting aside morphology, the term "syntactic doubling" may not be entirely appropriate in describing the portmanteau constructions in code-switching, since "syntactic doubling" may refer to some monolingual phenomena (Barbiers, 2008; Barbiers et al., 2008) which, as Hicks (2010) cogently argues, are very different in nature. The following are sampled from the phenomena discussed in the volume of Barbiers et al. (2008).

(24) **An** a   **han** joort hi
He has he   done it
"He has done it.
(Finland Swedish, Barbiers, 2008, p. 11)

(25) He **should can** go tomorrow.
(Scottish English, Barbiers, 2008, p. 16)

(26) Jan **kan** best schaatsen **kunnen**
Jan can best skate   can.INF
"It is perfectly possible that John is able to skate."
(Dutch, Barbiers, 2008, p. 17)

(27) Ek ken   **nie** daardie man **nie**
I   know not that   man not
"I don't know that man."
(Afrikaans, Biberauer, 2008, p. 104)

(28) **ä** ganz **ä** liebi   frau
a really a lovely wife
"a really lovely wife"
(Swiss German, Barbiers, 2008, p. 5)

(29) **Leer**,   ningún estudiante ha **leido** este libro
Read-INF no   student   has read this book
"As for reading, no student has read this book."
(Spanish, Barbiers, 2008, p. 16)

(30) **Um**   hvað eruð þíð   að tala **um**
about what are   you-PL to talk about
"What are you talking about?"
(Icelandic, Jónsson, 2008, p. 404)

Doubling of pronouns such as (24) is distinct from the portmanteau construction in the sense that the doubled pronouns have no complements. In (25), however, the two modal auxiliary verbs (i.e., "*should*" and "*can*") share a complement VP ("i.e., "*go tomorrow*"). In a sense the syntactic category of

---

[3]In fact, neither Hicks (2010) nor Stolz (1996) makes this explicit. Stolz (1996), however, does imply this as he glossed "−*sti*" with an abbreviation "COO," and elsewhere only the bound morphemes seem to be glossed with an abbreviation. With a view that conjunctions are rarely clause-internal, which appears to be the case for "−*sti*" in (23), I assume that "−*sti*" is a bound morpheme. At any rate, the point here is that it is possible to find doubling of synonymous elements involving a word and a bound morpheme.

[4]For instance, there are more or less the same number of VO or OV languages in the world (Dryer, 2013b), but cross-linguistically suffixes seem overwhelmingly more prominent than prefixes, and infixes are much less common (Dryer, 2013a).

modal verb, presumably a functional head in the I or T domain, is doubled, but the modals here are two different words of two different meanings. In portmanteau sentences, the doubled heads appear to be of the same word though realized in two different phonetic forms associated with two separate languages. In (26), the doubled modal auxiliary is of the same word ["*kan* (can)"], and in terms of surface order this example looks very similar to a portmanteau construction in which two instances of "*kan* (can)," supposedly an I or T head again, surround a complement VP ["*best schattsen* (best skate)"]. However, as shown in the English translation and explained by Barbiers (2008, p. 17), the two instances of "*kan*" convey quite different meanings; that is, the first one is epistemic and has scope over a proposition (i.e., *It is* **possible** *that PROPOSITION)* whereas the second one denotes the subject's ability (i.e., *John* **is able to** *skate*). In portmanteau sentences, the doubled heads appear to carry essentially the same meaning[5]. In (27), the doubled negation markers do seem to convey the same meaning, but, as Biberauer (2008) explains, only the first "*nie1*" is a NEG head merged in VP, whereas the second one, "*nie2,*" is really a Polarity Head above CP (that dominates the VP). The first "*nie1*" moves up to the specifier position of the Polarity phrase with VP, resulting in the "*nie1* VP *nie2*" sequence. This movement account does not extend to portmanteau sentences, if we assume that the doubled head (X_A and X_B) is of the *same* syntactic category (e.g., V, C, T, or P)[6].

In (28), the doubled indefinite determiners do seem to be the same word conveying the same meaning. Nonetheless, the first one, which is optional, is licensed by a degree or quantity expression [e.g., "*ganz* (really)," also see Kallulli and Rothmayr, 2008]. In portmanteau sentences, neither instance of the doubled heads seems to be licensed by an element other than its complement (i.e., the complement is obligatory in a portmanteau sentence). In (29), again, the doubled heads, which is a verb in this case, are the same word with the same meaning, but the first one ["*leer* (read)"] carries an intransitive reading whereas the second one ["*leido* (read)"] is transitive (Barbiers, 2008). In portmanteau constructions, both verbs are transitive and argument-sharing. In (30), the doubled prepositions "*um* (to)" do share the same complement [i.e., "*hvað* (what)"], but the second one is far away from the complement which has undergone wh-movement (Jónsson, 2008). In portmanteau sentences, both instances of the doubled head are both contiguous to their complement.

Having pointed out the differences between the portmanteau construction as a kind of "syntactic doubling" and "syntactic doubling" in monolingual phenomena, it would be fair to mention that in fact the names "portmanteau" (Nishimura, 1995, p. 157) and "palindromic switches" (Sankoff et al., 1990, p. 52) are not necessarily better descriptions of the code-switched construction being discussed. The term "portmanteau"

is supposed to refer to "blends" originally (e.g., "*smog*" that is blended from "*smoke*" and "*fog*")[7]. Portmanteau constructions in code-switching obviously do not refer to such lexical blends but they are more like "syntactic blends" (e.g., SVOV is blended from SVO and SOV). "Palindrome" denotes a series of linguistic items, including alphabets or words, which is the same whether reading forward or backward, such as "*madam*"[8]. Again, the portmanteau sentences are palindromic only in the sense of their syntactic sequence (i.e., X YP X). This paper adheres to the name of "portmanteau" because it is deemed a more popular one for the code-switched construction being examined.

This comparison with monolingual syntactic doubling and discussion of the names (i.e., "doubling" vs. "portmanteau" vs. "palindromic") is cursory and by no means comprehensive[9], but hopefully it serves to sharpen our focus on the so-called portmanteau construction in code-switching. In particular, we are dealing with cases of "lexical doubling" where the same word is realized into two synonymous but different phonetic forms. Of course, there is also "syntactic doubling" at the same time; that is, two words of the same category (i.e., X_A and X_B) appear in the sentence. However, the fact that these heads appear in positions adjacent to and on both sides of their shared complement seems better captured by the descriptors of "portmanteau" or "palindromic."

CONSTRAINTS ON THE PORTMANTEAU CONSTRUCTION

As commented by Sankoff et al. (1990, p. 52), "[p]alindromic switches, also known as portmanteau, copy translations, or mirror-image constructions, are widely attested but are inevitably found to occur rarely in quantitative studies." They continued, "Thus, these seem to constitute an occasional *ad hoc* production strategy rather than a systematic approach to bilingual sentence production" (Sankoff et al., 1990, p. 52).

Whereas, the portmanteau construction is indeed rare or unexpected in relation to not only monolingual phenomena but also code-switching, these data should not be automatically brushed aside as "periphery" (vs. "core," Chomsky, 1981) or performance data for which any attempt of systematic explanation is deemed futile. Crucially, portmanteau sentences have been attested in disparate speech communities and in different datasets, involving various language-pairs. It is at least a recurrent pattern in code-switching which is predictable in code-switching with typologically different languages. Additionally, in these language-pairs, portmanteau constructions are a general pattern which involves not only lexical verbs but also different kinds of heads (see above for a brief survey and see below for more details). Last but not least, it is clear that there are syntactic patterns or regularities that are amenable to more

---

[5]As for the portmanteau constructions, the authors, many of whom are bilingual in the two languages involved, appear to interpret the doubled elements as synonymous and equal in meaning, which is shown in the glosses and English translations (e.g., Nishimura, 1985, 1986; Park, 1990); see footnote 1 above.

[6]See below for more discussion about a movement analysis of portmanteau sentences.

---

[7]"Portmanteau" from Wikipedia (http://en.wikipedia.org/wiki/Portmanteau).

[8]"Palindrome" from Wikipedia (http://en.wikipedia.org/wiki/Palindrome).

[9]That is, there may be more deep-rooted similarities between monolingual syntactic doubling and portmanteau constructions which this section has not addressed and which are open to further research.

general explanation in terms of syntactic constraints, particularly the following.

(31) *Some heads do not double.*

The first regularity concerns the lack of data in which nouns, adjectives, modals, and conjunctions act as the doubled head in portmanteau constructions. It is not entirely clear whether the absence of these categories is due to empirical gaps (i.e., they are possible but they have not been attested) or some syntactic reasons. Worse still, grammaticality judgment, which potentially differentiates both scenarios, is not always reliable or consistent for code-switching since it may be affected by varying bilingual proficiency (MacSwan, 1999; Toribio, 2001), not to mention the irregularity of the portmanteau constructions under examination[10]. Based on the available data, tentatively speaking, the absence of modals or determiners may be just an empirical gap, if auxiliary verbs (e.g., (15), also supposed to be in I/T as modals) or some other functional heads in DP [e.g., (2), (21)] can be doubled in portmanteau constructions. On the other hand, there may be more deep-rooted reason underlying the absence of nouns (predicative) adjectives[11] and conjunctions.

Nouns are not found to partake in portmanteau constructions in language-pairs where the "noun complements" canonically appear on different sides of the head noun, such as Cantonese-English (Chan, 2008, 2015), Hindi-English (Pandit, 1986), or Tamil-English (Sankoff et al., 1990). In earlier frameworks such as X-Bar Theory (Jackendoff, 1977), nouns do take complements; for instance, a derived nominal or nominalization takes a DP complement [e.g., (32a)], see Chomsky, 1970), similar to the way in which its related verb takes an object [e.g., (32b)]. However, contrary to objects of transitive verbs which are obligatory, noun complements are grammatically optional [e.g., (32c)].

(32) a. the destruction of Rome
            N   P DP
         HEAD   COMPLEMENT

      b. The barbarians destroyed Rome.
                        V   DP
                     HEAD   COMPLEMENT

      c. The destruction caused a lot of casualties.
         N (HEAD)

Another difference between nouns and verbs is that nouns cannot take their complement directly. In Government-and-Binding Theory, this is because nouns lack case-assigning properties

(Chomsky, 1981). To introduce its complement, a case-assigner has to be introduced, such as a preposition in English or a "nominalizer" in Chinese languages which is most likely a functional head. Such nominalizers or genitive markers are attested in other languages where the "noun-complements" are prenominal, such as "ke" in Hindi (Pandit, 1986) or "uDaya" in Tamil (Sankoff et al., 1990).

(33) lo4-maa5 **ge3**   mit6-mong4
      Rome   NOM   destruction
      "The destruction of Rome"
                                  (Cantonese)

This case-based account, however, does not explain "*that*" which is required to introduce sentential complements of nouns (e.g., (34)—Haegeman and Guéron, 1999, p. 440); these sentential complements are not supposed to bear case.

(34) The news *(that) Peter has resigned bothered me.

            N   CP
         HEAD   COMPLEMENT

In view of the optionality of the so-called "noun-complements," Kayne (2009) proposes that they are in fact a variety of relative clauses, which are an adjunct rather than a complement. In other words, nouns actually do not take complements. If this is on the right track, then it is not surprising at all that nouns do not take part in portmanteau constructions, an integral condition for which is that a head merges with its complement and projects a phrase with the same label.

Not much is known about the case of adjectives. Attributive adjectives are standardly analyzed as an adjunct or a specifier of a functional head in the Cartographic Approach to syntax (e.g., Cinque, 2005) which does not take a complement. Some predicative adjectives do seem to license internal arguments but at least in English they do not take them directly; similar to the case of nouns, a preposition is called for to introduce a complement.

(35) The manager is open *(to) different suggestions.
            ADJ   P DP
         HEAD   COMPLEMENT

Pending confirmation from further research, it is plausible that at least in some languages (e.g., English) adjectives do not project to an Adjective Phrase with a complement either. If this were a more general phenomenon across languages, an adjective would not take part in portmanteau constructions, even though it might canonically appear on both sides of its internal argument [e.g., "*different suggestions*" in (35)] in the languages that a bilingual speaks. At any rate, there seems little existing data of code-switching which involve two languages in which predicative adjectives show different head-complement order.

Assuming that conjunctions are a functional head on a par with complementizers (C), determiners (D) and *do*-auxiliary verbs, which take part in the portmanteau construction, we expect to find conjunctions being doubled in a portmanteau

---

[10]Grammaticality judgment of code-switched sentences has been widely assumed to be affected by bilinguals' varying proficiency in their two languages. Toribio (2001) finds that a group of Spanish-English bilinguals of different proficiency levels show varying grammaticality judgments. Other researchers were prudent in choosing only competent bilinguals to be their subjects (e.g., MacSwan, 1999; González-Vilbazo and López, 2011, 2012). The point here is not to suggest that grammaticality judgment is never valid in code-switching studies, but it may be problematic to ask different groups of bilinguals (who speak various language-pairs) to give grammatical judgments of portmanteau sentences, including patterns attested in the data and other hypothetical ones.

[11]Attributive adjectives are standardly assumed to be an adjunct but not a head (Santorini and Mahootian, 1995), but some scholars do take attributive adjectives to be a head (Cantone and MacSwan, 2009).

construction too. Contrary to expectation, there are few instances of the portmanteau construction involving a doubled conjunction [except (23) above in which one conjunction is an affix attached to the subject noun]. One possible reason is that a conjunction rarely appears after the second conjunct clause (i.e., [XP CONJ YP] is possible but [XP YP CONJ] is much rarer)[12]. At any rate (as far as I am aware of) there is not any attested evidence of code-switching between a language that licenses an [XP CONJ YP] order and another that allows an [XP YP CONJ] order.

The absence of a sequence of [XP YP CONJ] is very much a logical consequence if we subscribe to Chomsky's (2013, p. 46) recent suggestion that CONJ does not merge with a conjunct clause (i.e., XP or YP) but a sequence of [XP YP]. Failure to label the phrase [XP YP] drives the movement of XP above CONJ, resulting in [XP CONJ X̶P̶ YP] which is labeled as an XP but not a CONJP. In such an account, CONJ is not a projecting head in the sense that it does not first-merge with its complement (e.g., YP) and project a phrase (i.e., *CONJ YP →[CONJP CONJ YP]).

(36) *Complements do not double.*

A second recurrent pattern is that it is the head that is doubled, but never (to the best of my knowledge) are there data in which the complement is doubled rather than the head. If this possibility sounds outlandish, we may be reminded that in the minimalist program all derivations are possible unless they are "crashed" for some reason (Chomsky, 1995; MacSwan, 1999). In other words, the impossibility of a [YP $X_{A/B}$ YP] sequence calls for an explanation.

The absence of an SOVO pattern may be explained by the classic theta-criterion in the earlier Government-and-Binding Theory[13].

(37) *The Theta-Criterion* (Chomsky, 1981, p. 36)
    a. Each theta role is assigned to one and only one argument.
    b. Each argument is assigned one and only one theta role.

In accordance with (37a), one object in an SOVO structure would not receive a theta/thematic role, hence the impossibility of such a sequence. On the other hand, although the subject and object in an SVOV structure apparently receive theta-role twice from the two reduplicated verbs, these reduplicated verbs are arguably the *same* word (see above) and hence also the *same* verb, and supposedly the subject and the object still receive one and *the same* role (e.g., Agent for Subject, Theme for Object, Recipient

for the Indirect Object, etc.)[14]. However, this explanation cannot account for other types of portmanteau constructions in which the doubled element are from other categories which do not assign theta-roles (e.g., copula verb or auxiliary verb).

(38) *Word order (i.e., head-initial vs. head-final) always follows the language of the head*

A third regularity is that the head from head-initial language (e.g., X_A) always remains head-initial whereas that from the head-final language (e.g., X_B) always stays head-final in the portmanteau sentences. Whereas, this sounds self-evident or merely descriptive, this regularity is not to be taken for granted particularly in SVOV structures, since in non-portmanteau, code-switched sentences a verb from a head-initial language can appear in head-final position, and in reverse a verb from a head-final language can also appear in head-initial position. The following are some examples from various typologically different language-pairs surveyed in Chan (2003, 2008, 2009).

V from OV language, VO order
(39) **pooTuruvaan**   *letter*
      put (3-SG-MASC-FUT) letter
      "He will write a letter."
                                    (Tamil-*English*, Sankoff et al., 1990, p. 79)

(40) I have to ***ttakē* my hand**
      I have to wash my hand
      "I have to wash my hand."
                                    (English-*Korean*, Choi, 1991, p. 889)

V from OV language, VO order
(41) i   ka   **i**   *rectify*
      3PL TAM 2SG rectify
      "They usually rectify you."
                                    (Mandinka-*English*, Haust and Dittmar, 1997, p. 88)

(42) want   on   Tex laat ons **daai (daardie)** *group join*
      because old Tex make 1PL DEM   group join
      "Because old Tex made us join that group."
                                    (Tsotsitaal-*English*, Slabbert and Myers-Scotton, 1996, p. 332)

The following two examples also show OV order with a verb from a VO language, but they involve the so-called "mixed compound verb" structure in which an auxiliary verb "*do*" appears in I/T (Chan, 2003, 2008)[15].

(43) kamalaa ne   hamaare ghar   par ***chicken taste*** kiyaa
      Kamla   ERG our   house at chicken taste did
      "Kamla tasted chicken at our house."
                                    (Hindi-*English*, Pandit, 1986, p. 106)

---

[12]Haspelmath (2007) points out that [XP YP CONJ] does exist in some languages, for example, *que* in Latin (Haspelmath, 2007, p. 8), but it is much rarer than an [XP CONJ YP] sequence cross-linguistically. Even in the exceptional case of the conjunctive morpheme in (23), it is more likely to be clause-initial in syntactic derivation but eventually gets affixed to the subject noun via lowering at PF (Embick and Noyer, 2001). The possibility that it is merged in clause-final position seems far more remote.

[13]Apparently, an SOVO structure would also violate the head-initial/head-final value parameterized for the language of the verb. However, the picture is more intricate in code-switching in which the language of the verb does not necessarily determine head-complement order (Chan, 2003, 2008, also see below).

[14]Even if the doubled verbs are seen as two verbs, it is not problematic to assume that the subject and object of portmanteau sentences receive two theta roles. In monolingual syntax, arguments in complex predicates or the subject in control sentences seem to receive two roles too, but from two different verbs or predicates (Ackema, 2014).

[15]See chapter 7 in Muysken (2000) more in-depth discussion of the mixed compound verbs.

(44) anta **car-ei**   *drive* paNNanum
　　 that car-ACC drive do + must
　　 "We must drive that car."

<div align="right">(Tamil-English, Sankoff et al., 1990, p. 80)</div>

Such "mismatch" between the language of the verb (i.e., VO or OV) and the order of the code-switched phrase, however, does not extend to other heads. From available data, the language of a functional head, including adpositions, always determines head-complement order in code-switching, either in portmanteau and non-portmanteau sentences (e.g., English preposition always remains prepositional and a Cantonese postposition always remains postpositional in a code-switched PP—Chan, 2015).

PORTMANTEAU CONSTRUCTIONS AND SYNTACTIC MODELS OF CODE-SWITCHING

Neither are the form of portmanteau constructions and the constraints on them captured by current syntactic models of code-switching (Hicks, 2010, 2012), including the Equivalence Constraint (Poplack, 1980), the Matrix Language Frame Model (Myers-Scotton, 1993, 2002; Myers-Scotton and Jake, 2009), the Bilingual Speech Model (Muysken, 2000, 2013) and the Null Theory (Mahootian, 1993; MacSwan, 1999, 2000; Chan, 2003, 2008).

As a classic that stimulated much subsequent work on the syntax of code-switching, the Equivalence Constraint (Poplack, 1980) prohibits code-switching at points where the surrounding words have divergent word orders in the participating languages (Poplack, 1980, p. 228; Sankoff and Poplack, 1981, p. 5–6). Accordingly, with switches between head and complement within a phrase (e.g., DP, VP, PP, or CP) whose word orders contrast in the participating languages (i.e., head-initial vs. head-final), portmanteau constructions violate the Equivalence Constraint (Poplack, 1980; Hicks, 2010, 2012)[16]. Looking at Tamil-English, which is one major source of the portmanteau data, Sankoff et al. (1990, p. 92) acknowledge the violation. However, they think that these constructions are a way to circumvent the Equivalence Constraint since the word orders of both languages are respected "as the lesser of evils" (Sankoff et al., 1990, p. 92). This idea is sensible, but counter-examples of the Equivalence Constraint in data other than portmanteau constructions (see Chan, 2003 for a survey) weaken the validity of the constraint and the feasibility of this suggestion.

If the Equivalence Constraint is unrealistically too narrow in confining code-switching to where two grammars overlap in bilingual competence (Poplack, 1980, p. 612, Figure 3; Sankoff and Poplack, 1981; Woolford, 1983), the Matrix Language Frame Model (Myers-Scotton, 1993, 2002) is certainly broader in empirical scope, but nonetheless there is a baseline. That is, the

grammar of the more dominant language, that is, the Matrix Language (or ML), has to be observed. This is supposed to be the case since the Matrix Language alone constructs the "frame" of a code-switched sentence (via *the Uniform Structure Principle*—Myers-Scotton, 2002, p. 8–9). There are two ways in which this is accomplished. One, ML sets the word order of a code-switched sentence via *the Morpheme Order Principle*, and, two, ML provides the "system morphemes," mostly function words or bound morphemes, via *the System Morpheme Principle*. The less dominant language, namely, the Embedded Language, only contributes content words (or "content morphemes") or phrases (i.e., "EL islands" which nonetheless are formed in EL grammar by virtue of the *EL Island Principle*) to be inserted into the frame. Being another paradigm in the syntax of code-switching, there has been much follow-up discussion and extension of the original model (Myers-Scotton, 2002, 2006; Myers-Scotton and Jake, 2009), in particular, the appended "4M" Model. It proposes a more fine-grained classification of system morphemes so that the "early system morphemes" (such as determiners or plural suffixes) may be activated from EL but the "bridge" morphemes (e.g., the non-theta assigning preposition "*of* ") or the "outsider" morphemes (e.g., agreement markers) are rarely accessed from EL (Myers-Scotton, 2002, 2006; Myers-Scotton and Jake, 2009).

Concerning the portmanteau construction, it is sufficient to note that the juxtaposition of word orders from both languages in a code-switched sentence makes it impossible to designate the Matrix Language and hence also the Embedded Language in that sentence, since the design of the model requires that one participating language has to be the ML and the other be EL (in accordance with *the Asymmetry Principle*, see Myers-Scotton, 2002, p. 9). In other words, the Morpheme Order Principle has to be violated by the portmanteau constructions; to be more concrete, referring back to (1), if language A were the ML, the order of YP X_B would be against the word order of the ML; if language B were the ML, the order of X_A YP would contradict the word order of ML. Also challenged are the Asymmetry Principle (which dictates that one language is ML and the other is EL) and the Uniform Structure Principle (which stipulates that ML alone contributes to the structure of a code-switched constituent—Myers-Scotton, 2002, p. 8–9)[17].

---

[16]Note that portmanteau constructions do not necessarily violate the Equivalence Constraint if the switch takes place within YP (e.g., X_A [YP] X_B) the internal order of which is shared between language A and language B (e.g., [$_{DP}$ D ADJ N]). However, in most existing data [i.e., (2), (3), (6)–(18)], the switch falls either between X_A and YP or between YP and X_B, thus contravening the constraint.

[17]Apparently not having touched upon the portmanteau constructions so far, Myers-Scotton (1993, 2002) did tackle examples of morphological doubling [e.g., doubling of the plural affix in (22)]. The explanation is that, as a bound morpheme is retrieved from ML [e.g., the Shona plural affix "*ma*" in (22)], its counterpart in EL [e.g., the English affix "−*s*" in (22)] is also co-activated, resulting in morphological doubling. Even though this apparently violates the System Morpheme Principle (i.e., a bound morpheme, supposedly a "system morpheme," is drawn from EL), the grammar of ML is still respected with the doubled morpheme from ML. The idea was formalized as *the Double Morphology Principle* in Myers-Scotton (1993, p. 133). Under the "4M" Model, the appearance of an EL system morpheme is even less of a problem (Myers-Scotton, 2002, p. 91–93) so long as it is not an "outsider" system morpheme. Putting aside the desirability of such reasoning (i.e., more machinery is invoked to deal with apparent counter-examples, see Chan, 2003, 2009) it seems difficult to apply similar argumentation to the portmanteau construction (i.e., a head from EL is allowed as long as the corresponding head from ML is lexicalized). Crucially, the doubling of heads involves word order, hinging upon the Morpheme Order Principle. Additionally, many examples do involve a switch to longer elements or more than one instance of syntactic doubling

In a theoretical perspective, the Matrix Language Frame Model may be too "heavy" in invoking a "grammar" specific to code-switching (Chan, 2003, 2008, 2009). In this light, greater theoretical and cognitive economy is achieved in attempts that subsume recurrent patterns in code-switching into general constraints independently proposed for monolingual phenomena, an early one being the Government Constraint (DiSciullo et al., 1986). The idea that code-switching and monolingual sentences are governed by the same linguistic constraints and mechanisms has eventually been dubbed "the Null Theory" since Mahootian (1993), inspiring later works (MacSwan, 1999, 2000; Chan, 2003, 2008, also see MacSwan, 2014). Strictly speaking, the Null Theory is *not* one coherent theory but more of a theoretical position, and studies that claim to follow "the Null Theory" may make different empirical predictions because of the various syntactic theories or constraints they appeal to respectively [e.g., Tree-Adjoining Grammar for Mahootian (1993), The Principles-and-Parameters Framework for Chan (2003, 2008), or the Minimalist Program for MacSwan (1999, 2000) and the papers in MacSwan (2014), see Chan, 2009 for a summary]. Not surprisingly, the term "Null Theory" is seldom mentioned in more update work in a similar vein that does not presume specific constraints on code-switching (e.g., González-Vilbazo and López, 2011, 2012; Shim, 2013, or the papers in MacSwan, 2014). No matter what specific theory or version of a theory it is, portmanteau sentences are problematic for the Null Theory, because they are radically different from monolingual phenomena, a construction that presumably arises out of language contact and hence is specific to code-switching (Chan, 2009, also refer back to the above section on differences between the portmanteau construction and monolingual doubling). This is not to suggest that theories of monolingual syntax can never be extended to the syntax of portmanteau constructions, but the "bilingual element" that sets apart monolingual sentences and the portmanteau ones has to be identified and captured in any satisfactory account of the latter.

Though put forward by a veteran in generative linguistics, the Bilingual Speech Model (Muysken, 2000, 2013) presents a rather different vision from that of the more recentstudies which continue to explore possible constraints on a specific dataset or language-pair with reference to facets of the Minimalist framework (e.g., González-Vilbazo and López, 2011, 2012; Shim, 2013). More comprehensive and "variationist" in outlook, the Bilingual Speech Model (Muysken, 2000, 2013) envisages different strategies with which bilinguals or bilingual communities engage in "code-mixing" or intra-sentential code-switching. *Alternation* refers to a total switch to another language in lexis and grammar, whereas by *insertion* a word or a phrase is inserted to a sentence framed by the Matrix Language. *Congruent lexicalization* is a third strategy where a code-switched sentence has a structure shared between the two participating languages and so words may be drawn from

_____

[e.g., (2), (7)–(9) so that it does look virtually impossible to assign an ML and an EL for that code-switched sentence.

either language anywhere in the sentence without constraint. The fourth one, namely, *backflagging*, is the latest addition (Muysken, 2013), in which a bilingual speaker uses some elements of his/her heritage language even though he or she has shifted to a new language. In this framework, Muysken (2000, p. 104–105) does describe the portmanteau constructions, which he calls "doubling," as *alternation*. This proposal is seconded by Takagi (2007) who renames the portmanteau construction as "symmetrical sentences" with reference to her dataset of Japanese-English code-switching produced by bilingual children. Namba (2012a,b) follows suit, but elaborates that portmanteau constructions are better treated as alternation and triggering (Clyne, 1987). For instance, in English-Japanese code-switching, a switch from an English verb to a Japanese object triggers Japanese grammar and eventually the doubling of a verb in Japanese [e.g., (2)]. However, alternation, which implies a long element after a switch, is problematic in capturing cases where there is only one word after a switch [e.g., (6), (11), (18), (19)], since the single switched word does not clearly show that the sentence switches to another "grammar." Even though the speaker code-switches to a longer fragment, alternation may still be awkward in describing examples where there are further switches after a speaker has alternated once [e.g., (12), (16), (17), (20)], since alternation denotes a "total" or "complete" switch in lexis and grammar [as illustrated in (2), (3), (7)–(10)]. Defined as extensive code-switching in a structure shared by both languages, congruent lexicalization (Muysken, 2000) does not apply to the portmanteau construction which involves contrasting word orders from both languages. Defined as occasional switching to a heritage language that the bilingual speakers seldom use in daily life, backflagging (Muysken, 2013) does not seem to apply to the portmanteau construction either, since the bilingual speakers who produce them do seem to be using both languages actively (if not equally actively) in their life.

The optionality of portmanteau constructions (i.e., non-portmanteau constructions, e.g., SVO or SOV, are also found in code-switching with typologically different languages) appears to invite an account along the lines of Optimality Theory. In the literature, however, there are not many studies of code-switching employing an Optimality-Theoretic framework. Among these few studies, Bhatt (1997, 2014) proposes that there are different constraint-rankings for code-switching involving different language-pairs. It is hence unclear how he may account for variant patterns of code-switching involving the same language-pair, such as portmanteau vs. non-portmanteau constructions. Focusing on Cantonese-English code-switching in a PP, Leung (2001) suggests a constraint-ranking which governs possible output of constructions. In brief, he concludes that the portmanteau construction (i.e., a PP involving an English preposition and a Cantonese postposition) and the non-portmanteau one (i.e., PP containing only the English preposition) are both allowed but other possible structures are forbidden. While the account successfully captures the empirical facts, the idiosyncrasy of the portmanteau pattern and its emergence remain opaque.

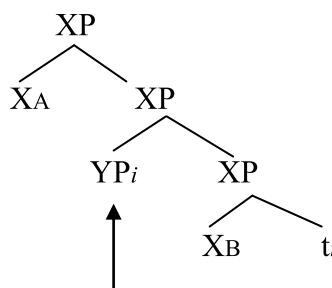## PORTMANTEAU CONSTRUCTIONS, PHRASE STRUCTURE AND LINEARIZATION

To account for the portmanteau construction, a fundamental issue that needs to be addressed is what kind of structure it may have. Apparently, the phonetic realization of two heads sharing the same complement suggests that the phrase may be ternary-branching rather than binary branching, which has been a mainstream assumption in generative grammar, particularly with reference to the Antisymmetry thesis proposed by Kayne (1994, 2004, 2005, 2009)[18].

(45)

$$XP \rightarrow X_A \quad YP \quad X_B$$
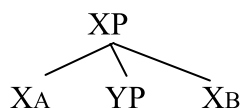
One could argue that the portmanteau phrase XP may be derived as follows in accordance with Antisymmetry and binary branching; that is, YP follows X_B and then it moves up before [YP X_B] merges with X_A.

(46)

The derivation in (46) is of course very much simplified. Firstly, in all data of portmanteau constructions involving the verb (e.g., (2), (3), (6)–(14) above), the doubled verb is the *main verb* of the sentence inflected for tense and agreement. Accordingly, the derivation involves the doubling of not only V but also T, as sketched below:

(47)
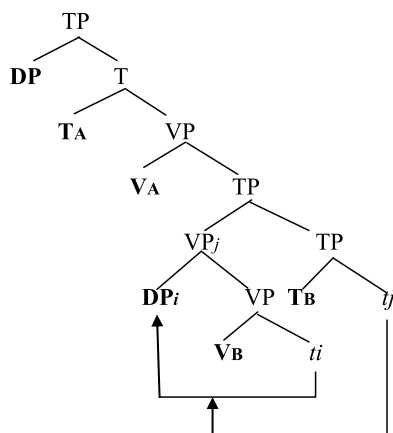
(A = a head-initial language; B = a head-final language)

In case the doubled head is a ditransitive verb [e.g., (6)] or a "saying" verb which takes a DP and a clausal object [e.g., (9)], more layers of vP shells (Chomsky, 1995) have to be invoked between T and V, resulting in even more derivational steps than in (47)[19].
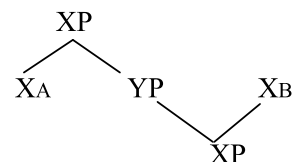
Several questions then arise. Are these derivations absolutely necessary? Is there a more economical way of capturing the portmanteau construction? Additionally, within the minimalist architecture of grammar, an outstanding question is why two words of identical meaning are simultaneously introduced to the Numeration (Chomsky, 1995). No matter what the answers to these questions are, there is a sense that they may well lie outside "syntax proper," even though the portmanteau constructions, as illustrated and argued in this paper, show recurrent syntactic patterns that are subject to structural constraints.

Difficulties to account for the portmanteau constructions in generative grammar suggest that these structures may be better handled by alternative models of grammar whose assumptions are radically different, for instance, functionalist theories such as Cognitive Grammar (Taylor, 2002; Langacker, 2008) or Radical Construction Grammar (Croft, 2001). However, this does not appear to be the case. Briefly speaking, these grammars focus on the meaning of constructions which are not seen as being built up by derivations, and the language faculty is not autonomous but connected to other cognitive functions or faculties. In these frameworks, the portmanteau sentences would convey some meanings that are distinct from those of their non-portmanteau counterparts. Nonetheless, this is far from clear in the data and their descriptions in the relevant literature. A related issue is that, if portmanteau constructions do not convey some additional or different meaning, these sentences would violate the principle of economy (Haiman, 1983, 1985; Croft, 2002; Chan, 2009).

TOWARD A MIXED ACCOUNT OF SYNTAX AND PROCESSING

In an innovative account, Hicks (2010, 2012) suggests that a bilingual accesses two sets of syntactic information and projects a dual structure for the portmanteau constructions, borrowing Sadock's (1983, 1991) Autolexical Syntax. A portmanteau phrase would have a structure as (48) with an upper layer and a lower layer.
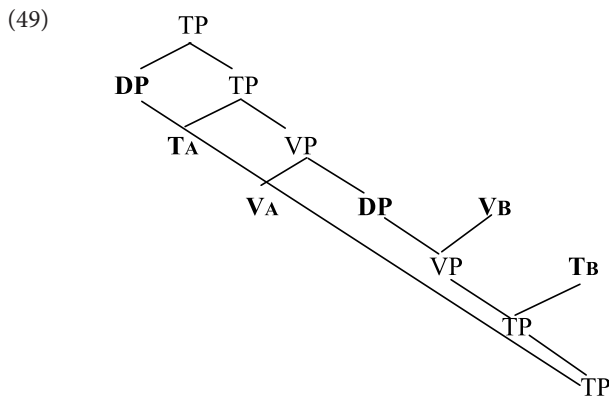
(48)

More elaborately, an SVOV sequence would have the following structure under this account.

---

[18]Ternary or "flat" structures are allowed in Simpler Syntax (Culicover and Jackendoff, 2005), however.

[19]In current phase-based theory (Chomsky, 2013; Citko, 2014), the light v is always introduced into the derivation of a sentence.

(49)

(Adapted from Diagram 3 in Hicks, 2012, p. 52).

The idea that bilinguals have access to two sets of syntactic information is intuitively convincing and uncontroversial. There is much psycholinguistic evidence that when a bilingual speaker processes or produces one language (i.e., the "target" language), the other language is also accessed (i.e., the "non-target" language, see Wu and Thierry, 2010 for an overview). However, co-access of syntactic information itself is too general a factor to explain the constraints on portmanteau constructions and the optionality of them. The constraints suggest that only certain types of syntactic information are responsible for the production of portmanteau constructions, and the optionality implies that there is a mechanism to filter out one language or one set of syntactic information, hence leading to non-portmanteau constructions in output.

The first issue is quite straightforward. It is empirically clear that portmanteau constructions emerge in language-pairs in which head-complement order is different for a particular phrase (e.g., VP, TP, CP, PP, DP). Additionally, it looks very plausible that only *projecting heads* take part in portmanteau constructions. Heads which arguably do not project (i.e., (first-)merging its complement), such as nouns, adjectives or conjunctions, do not take part in portmanteau constructions. Crucially, doubling of heads is highly related to *projection* (i.e., X merges with YP and results in an XP, i.e., [XP X YP]—Chomsky, 2013). Complements (e.g., YP) do not project (i.e., *[YP X YP]), and thus they do not double (i.e., *[YP YP X YP]).

The second issue calls for a distinction between *access* to syntactic information and *activation* of syntactic information (i.e., the syntactic information is processed, leading to an output, a phrase or a sentence). Presumably, bilinguals always have access to information of both languages, but they do not always activate both sets of information, for instance, when they are using only one language. This is consistent with the model of Language Modes (Grosjean, 2008, etc.) in which bilinguals may activate just one language with the other deactivated (i.e., the Monolingual Mode), or they may activate both (i.e., the Bilingual Mode). Level of activation nonetheless is relative and hence the Monolingual Mode and the Bilingual Mode are two ends of a continuum, and the mode of a bilingual is affected by many performance factors such as the context of speaking, the other participants, his or her language proficiency, etc. An alternative conception is suggested in Green and Li's (2014) model, in which two languages are always active in the mind of bilinguals

who engage in code-switching, but only some information is selected for output and other information is inhibited through some "control" mechanism (Green, 1986). In what Green and Li (2014) call "Competitive Control," a bilingual may speak in one language only, and information of the other language (e.g., words and morphosyntactic rules) is inhibited. In other contexts, a bilingual may engage in extensive code-switching, exercising less inhibition and allowing information of both languages to be processed further for output; Green and Li (2014) describe this cognitive process as "Open Control."

In Green and Li (2014) model, types of information about a language include word forms and syntactic constructions (or, more technically "Combinatorial Nodes"), and they are all linked in a network. Assuming that head-initial and head-final orders are two of such "combinatorial nodes" (i.e., [X YP] and [YP X] respectively), bilinguals of typologically different languages always have access to both sets of head-complement orders. However, portmanteau constructions (e.g., SVOV) arise when bilinguals do not inhibit either set in output. When they let only one set of order enter output (with the other inhibited), non-portmanteau constructions (e.g., SVO or SOV) would be the result.
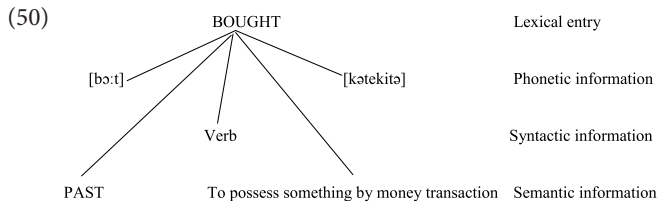
This way of capturing linearization in the production process may seem uneconomical and a radical departure from more mainstream accounts which envisage a more direct mapping between syntactic structure (i.e., a syntactic tree) and word order (e.g., Kayne, 1994, 2004, 2005). However, it is not inconsistent with the more recent view that linearization is a process at the Sensory-Motor (SM) interface (alternatively known as the interface of Phonetic Form/PF), and that syntactic structures are not specified for linear order in derivations (Chomsky, 1995, 2005, 2013; Kremers, 2009, 2012). With reference to the portmanteau construction, locating linearization in the production process avoids complexity of structure and derivations which plagues an Antisymmetry approach and to some extent Hicks' (2010, 2012) dual-structure account[20].

Let us further assume that a lexical item that enters a Lexical Array or Numeration and then syntactic computations is actually a bundle of related information about a word, including meaning, syntactic information (e.g., word class) and morpho-phonological information, largely equivalent to what is called a *lexical entry* in the psycholinguistic literature (e.g., Levelt, 1989)[21]. It seems that Chomsky (1995, 2005, 2013) has not rigorously defined what he meant by a lexical item except the comment that it must provide a label so that the Lexical Array will recognize it as head in projection (Chomsky, 2013, p. 43), with "label" presumably referring to the word's syntactic category. At any rate, "a lexical item" cannot just refer to a "word" in a
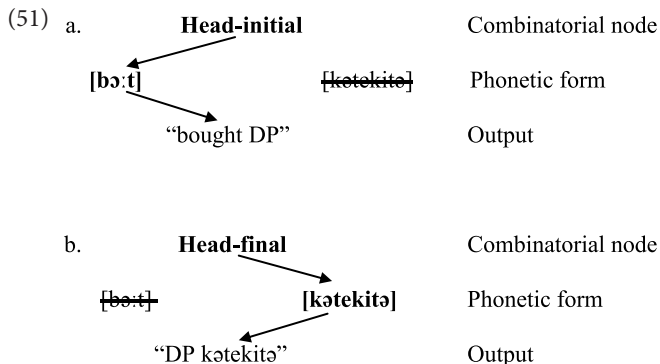
---

[20]There may be two more limitations of Hicks' (2010, 2012) dual-structure account. Firstly, it is not apparent that the doubled heads are of one word (see text above), since they are dominated by different layers of structure. This also leads to the second drawback; that is, the dual-structure does not show that a [X YP X] sequence is intuitively one phrase. However, the dual-structure account may well capture cases in which the doubled heads are two different words with different meanings [i.e., (4) and (5)].

[21]Whereas "lemma" seems a more familiar term used to refer to information about a word, according to Levelt (1989), it does not include the morpho-phonological form of a word, and hence I use "lexical entry" instead in this paper.

conventional sense as a pairing of a phonetic or written form and a meaning, since it is supposed to enter syntactic computations when it has not yet been transferred (for pronunciation or writing) in the minimalist architecture of grammar. In sum, it is not inconceivable that a lexical item is a bundle of connected information about a word. Furthermore, in bilinguals' lexicon, a lexical entry consists of two phonetic forms[22] . With reference to example (2) above, this conception of "a lexical item" or a lexical entry for a bilingual lexicon is sketched below in (50):

(50)

BOUGHT		Lexical entry
[bɔːt] [kətekitə]		Phonetic information
Verb		Syntactic information
PAST To possess something by money transaction		Semantic information

When the bilingual speaks English, the head-initial word order is selected, and so is the phonetic form [bɔːt]. The speaker also has access to the corresponding form [kətekitə], but this information is inhibited [i.e., (51a)]. Conversely, when speaking Japanese, the head-final order is selected, calling for the form [kətekitə]. At the same time the speaker has to inhibit the corresponding form [bɔːt] [i.e., (51b)].

(51)	a.

Head-initial	Combinatorial node
[bɔːt] ~~[kətekitə]~~	Phonetic form
"bought DP"	Output

	b.

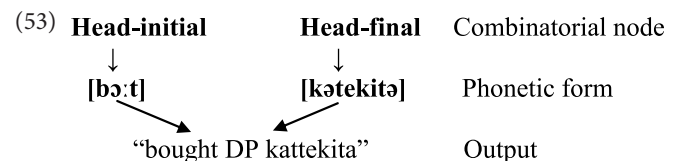|	
Head-final	Combinatorial node
~~[bɔːt]~~ **[kətekitə]**	Phonetic form
"DP kətekitə"	Output

It is reasonable to assume further that [bɔːt] is more often or strongly associated with the head-initial order whereas [kətekitə] is more strongly associated with the head-final order, and the two phonetic forms are linked between themselves by virtue of the fact that they are synonymous and belong to the same lexical entry [i.e., (52)].

(52)

BOUGHT	Lexical entry
[bɔːt] ———— [kətekitə]	Phonetic form
Head-initial Head-final	Combinatorial node
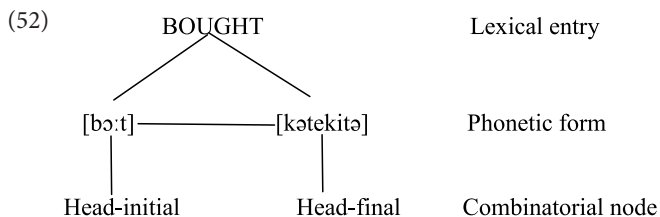
In a code-switching context, a bilingual may let both phonetic forms and both sets of word orders enter output without

---

[22]There is some psycholinguistic evidence that the two cross-linguistic synonyms or translations are both highly co-activated when a concept or *lemma* is activated (Costa et al., 2000; Runnqvist et al., 2013). See footnote 1 above.

inhibiting either. Here comes a very important issue. Are portmanteau constructions triggered by activation of both word orders (or "combinatorial nodes") or that of both phonetic forms? The former is much more plausible, if we assume that the bilingual mind is organized in the same way irrespective of the languages a bilingual speaks. That is, in case bilinguals speak both head-initial or both head-final languages, the activation of both phonetic forms would lead to sequences of SVVO or SOVV (or XXYP/YPXX when the doubled head is not a verb). Judging from the absence of these sequences (until they are documented in future), it appears that bilinguals do not usually activate both phonetic forms, even though this is actually possible under the Bilingual Mode (Grosjean, 2008). Therefore, portmanteau constructions are more likely to be motivated by the activation of both word orders which in turn call for the two corresponding phonetic forms.

(53)

Head-initial	**Head-final**	Combinatorial node
↓	↓	
[bɔːt]	**[kətekitə]**	Phonetic form
"bought DP kattekita"		Output

All the processes described in (51) and (53) are supposed to take place at the Sensori-Motor (SM) interface, the place where words are put in linear order and instructions are sent to the vocal organs to pronounce the words (it was called Phonetic Form (PF) in Chomsky's (1981, 1995) earlier works, largely equivalent to the stage of "planning" in language production models (Green and Li, 2014); in other words, the syntactic structure underlying the sequences of VO, OV (i.e., the non-portmanteau constructions) and VOV (i.e., the portmanteau constructions) is actually the same VP with the relative order of V and object DP unspecified.

(54)

VP	
V DP	
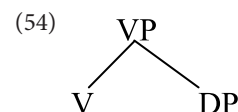
Generalizing this to portmanteau constructions where the doubled heads may not be a verb, the underlying structure of them is not exactly $[X_A \ YP \ X_B]$ as represented in (1) but simply an XP, with order between X and YP unspecified. Duplication of X_A and X_B arises in the Sensori-Motor interface as a bilingual activates both sets of head-complement order and realizes them with synonymous forms in two languages. Despite different phonetics these two forms are actually the *same* word belonging to the *same* lexical entry.
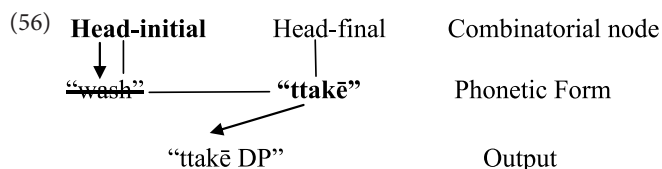
(55)

XP	
X YP	
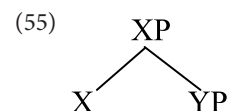
REMAINING ISSUES

There are a number of residual issues to be tackled. Firstly, the proposal so far has not fully explained the empirical fact that the language of the head always determines head-complement
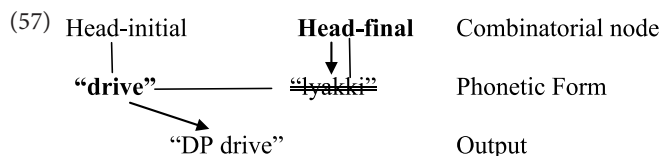
order in portmanteau constructions, especially when we consider that in non-portmanteau code-switched sentences the language of the verbs does not always determine head-complement order [see (39)–(44) above]. On the other hand, the language of functional heads (including adpositions) does seem to determine head-complement order in portmanteau and non-portmanteau sentences (Chan, 2003, 2008, 2015). The problem here is not so much about the portmanteau construction itself which is in a way explained by the proposal. The issue is really about how to explain verbs whose linear order does not match its "language" [i.e., a verb from a VO language appears in OV order or a verb from an OV language appears in VO order, e.g., (39)–(44)]. In addition, how do we account for the asymmetry between the verb and the other categories (i.e., C, I/T, D, also P tentatively) the language of which always determines head-complement order in portmanteau or non-portmanteau contexts?

Concerning the former issue, a recent syntactic account suggests that the properties of VP, including VO/OV order (in code-switching or in "pure" languages alike), are dependent on the feature composition of v (González-Vilbazo and López, 2011, 2012) and probably another functional head Asp(ect) between vP and VP (Shim, 2013). Putting aside how these models can be modified to accommodate the portmanteau construction (i.e., they do not specifically aim to explain the portmanteau construction), here we attempt to extend the psycholinguistic/production approach outlined above. Recall from (51) above that when the head-initial order is selected, the default case is that a verb associated with a VO language is also selected. This phonetic form, however, can be inhibited. As the processor is fast looking for a "substitute" for that form to produce the syntactic construction, the corresponding word form associated with an OV language is selected for output [i.e., (56)][23].

(56) **Head-initial**   Head-final   Combinatorial node

~~"wash"~~ ——————— **"ttakē"**   Phonetic Form

"ttakē DP"   Output

[Referring back to the English-Korean example in (40)].

Reversely, when a head-final order is selected, the default case is to activate a verb from an OV language, but this word form can be inhibited so that the corresponding word form from a VO language is selected for output [e.g., (57)].

(57) Head-initial   **Head-final**   Combinatorial node

"drive"———————— ~~"lyakki"~~   Phonetic Form

"DP drive"   Output

[Referring back to the Tamil-English example in (44)].

[23]This account does not address the issue of why a verb from a VO language is often accompanied by a light verb see (43) and (44) when it appears in OV order. See alternative accounts in González-Vilbazo and López (2011, 2012) and Shim (2013).

Under a psycholinguistic approach pursued here, the reason why a form is inhibited is prompted by performance factors including processing (e.g., the corresponding word form associated with another language has been more active in the context of speaking or "triggered" by a related form—Clyne, 1987, etc.) or pragmatics (e.g., that word form is deemed more appropriate in the context, i.e., the "mot juste," see Footnote 1).

These patterns [e.g., (39)–(44)] may well arise in a mental state between Competitive Control (Green and Li, 2014), where one language is selected and the other is inhibited, and Open Control (Green and Li, 2014), where information of both languages is allowed to enter output. In other words, a bilingual is speaking a selected language and yet the non-selected language is not completely blocked, and so some elements of the non-selected language may be selected for output. This is a state which Green and Li (2014) call Co-operative/Coupled Control.

In the case of functional heads, we may conjecture that their phonetic form is strongly associated with a combinatorial node and so it cannot be inhibited. Consequently, the mismatch between the language of a head and head-complement order [e.g., (56) and (57)] is not possible.

When both head-initial and head-final orders are activated, as in the case of portmanteau constructions, a bilingual presumably exercises less inhibition of information from both languages, that is, a state which is described as Open Control (Green and Li, 2014). Accordingly, both phonetic forms are activated without suppression of any one of them and they will go into their default position; for instance, the verb form associated with a VO language always goes into its default pre-nominal position and likewise its corresponding form associated with an OV language always appears post-nominally. The condition for the processor to find a "substitute" [i.e., the default word form associated with a combinatorial node is inhibited, as in (56) and (57)] does not exist anymore.

Now we turn to syntactic issues. The current proposal suggests that projection of a phrase triggers transfer and then linearization; however, there are two alternative scenarios as to the timing or the transfer. That is, either transfer is kickstarted as soon as a projecting head (first-)merges its complement along the lines of Kremers (2009), or it proceeds in *phases* in which a sentence is spelt out successively in vP and CP (Chomsky, 2005, 2013; Citko, 2014; also adopted in González-Vilbazo and López, 2011, 2012 and Shim, 2013 for code-switching). It is the standard phase theory which seems to provide a more unified explanation of portmanteau constructions involving different categories of a reduplicated head. More precisely, the "immediate linearization" approach can apparently explain portmanteau CPs [C being doubled, e.g., (8)], IP/TPs [I/T being doubled, e.g., (15)] and PPs [P being doubled, e.g., (17)], DPs [D being doubled, e.g., (21)], those instances in which verbs are doubled, which appear to be more common, suggest that linearization is procrastinated before it is transferred to the SM interface for linearization. That is, as these doubled verbs [e.g., "*bought*" in (2)] are morphologically inflected, they are supposed to move up to higher functional heads, for instance, the v head which selects VP (González-Vilbazo and López, 2011, 2012; Chomsky, 2013). The reason why the verb, contrary to other kinds of heads,

has to undergo further derivations is partly morphological and partly semantic (i.e., a verb takes up morphological marking to encode information such as tense, aspect and agreement). On the other hand, even though the other kinds of heads involved in portmanteau constructions do not seem to undergo further derivation or movement, it is not necessary that they must be transferred and linearized as soon as they project a phrase with their complement.

Thirdly, the copula verb may be doubled in the portmanteau construction, but it is not unanimously agreed that the copula projects a VP or copula phrase with its complement. In recent works, a copula verb merges with a small clause [XP YP] and XP raises eventually (e.g., [XP COP [X̶P̶ YP]]; Moro, 2010; Chomsky, 2013). There is however an alternative account in which the copula verb does project a phrase as a Relator head and merges with a predicate phrase as its complement (den Dikken, 2006). This account appears to be more consistent with this account of the portmanteau construction.

CONCLUSIONS

This paper proposes a combined syntactic and psycholinguistic account of portmanteau constructions in code-switching. The syntax side of the account crucially hinges upon the minimalist view that order is an interface phenomenon but syntactic structures (Chomsky, 1995, 2005, 2013; Kremers, 2009, 2012), at least those of a phrase in which a head

merges a complement, are not specified for order. One other assumption needed is that a lexical item which enters into a Lexical Array and eventually syntactic derivations is actually a "lexical entry" which is a bundle of various kinds of information about a word. In the case of bilinguals, this lexical item also contains information of a word in two languages. The psycholinguistic side of the account relies on Green and Li's (2014) model of Cognitive Processes of Control in which bilinguals may select one language for output and inhibit another, or they may let information of both languages be processed further for output. Crucially, projection of a phrase will lead to linearization, and a bilingual may co-activate and process both word orders (i.e., head-initial and head-final) if he or she speaks a head-initial and a head-final language. Whereas there is much work to be done to further clarify a number of issues pertaining to the account (in particular, whether the activation of both word orders is intentional or due to lapse of inhibitory control), this paper discusses an interesting case in which a limited set of performance data of a language-contact phenomenon, that is, the portmanteau construction, could lend empirical support to the ideas that a syntactic object is order-less and linearization is a process at the Sensori-Motor interface. At any rate, it is hoped that this work, despite all its limitations and stipulations, will raise more scholarly interest in the portmanteau construction and related issues, and stimulate more research on these topics.

REFERENCES

Ackema, P. (2014). "The syntax-lexicon interface," in *The Routledge Handbook of Syntax*, eds A. Carnie, D. Siddiqi, and Y. Sata (London; New York, NY: Routledge), 322–344.

Azuma, S. (1993). "Word Order vs. Word Class: portmanteau sentences in bilinguals," in *Japanese/Korean Linguistics*, Vol. 2, ed P. M. Clancy (Chicago, IL: CSLI Publications), 193–204.

Azuma, S. (1997). Lexical categories and code-switching: a study of Japanese/English code-switching in Japan. *J. Assoc. Teach. Jpn.* 31, 1–24.

Azuma, S. (2001). "Functional categories and codeswitching in Japanese/English," in *Codeswitching Worldwide II*, ed R. Jacobson (Berlin: Walter de Gruyter), 91–103.

Backus, A. (1996). *Two in One—Bilingual Speech of Turkish Immigrants in the Netherlands*. Tilburg: Tilburg University Press.

Backus, A. (2003). Units in code-switching: evidence for multimorphemic elements in the lexicon. *Linguistics* 41, 83–132. doi: 10.1515/ling.2003.005

Barbiers, S. (2008). "Introduction," in *Syntax and Semantics 36: Microvariation in Syntactic Doubling*, eds S. Barbiers, O. Koeneman, M. Lekakou, and M. van der Ham (Bingley: Emerald), 1–34.

Barbiers, S., Koeneman, O., Lekakou, M., and van der Ham, M. (eds.). (2008). *Syntax and Semantics 36: Microvariation in Syntactic Doubling*. Bingley: Emerald.

Bhatt, R. M. (1997). Code-switching, constraints, and optimal grammars. *Lingua* 102, 223–251.

Bhatt, R. M. (2014). "Argument licensing in optimal switches," in *Grammatical Theory and Bilingual Codeswitchig*, ed M. MacSwan (Cambridge, MA: MIT Press), 135–158.

Biberauer, T. (2008). "Doubling vs. omission in Afrikaans negation structures," in *Syntax and Semantics 36: Microvariation in Syntactic Doubling*, eds S. Barbiers, O. Koeneman, M. Lekakou, and M. van der Ham (Bingley: Emerald), 103–140.

Cantone, K., and MacSwan, J. (2009). "Adjectives and word order," in *Multidisciplinary Approaches to Code Switching*, eds L. Isurin, D. Winford, and K. de Bot (Amsterdam; Philadelphia, PA: John Benjamins), 243–277.

Chan, B. H.-S. (1998). "How does Cantonese-English code-mixing work?" in *Language in Hong Kong at Century's End*, ed M. C. Pennington (Hong Kong: Hong Kong University Press), 191–216.

Chan, B. H.-S. (2003). *Aspects of the Syntax, Pragmatics and Production of Code-switching—Cantonese and English*. New York, NY: Peter Lang.

Chan, B. H.-S. (2008). Code-switching, word order and the lexical/functional category distinction. *Lingua* 118, 777–809. doi: 10.1016/j.lingua.2007.05.004

Chan, B. H.-S. (2009). "Code-switching with typologically distinct languages," in *The Cambridge Handbook of Linguistic Code-Switching*, eds B. E. Bullock and A. J. Toribio (Cambridge: Cambridge University Press), 182–198.

Chan, B. H.-S. (2015). *Single English prepositions in Hong Kong Cantonese: Structural Borrowing and Convergence*. Macao: Department of English, The University of Macau.

Choi, J. (1991). Korean-English code-switching: Switch-alpha and Linguistic constraints. *Linguistics* 29, 877–902. doi: 10.1515/ling.1991.29.5.877

Chomsky, N. (1970). "Remarks on nominalization," in *Readings in English Transformational Grammar*, eds R. A. Jacob and P. S. Rosumbaum (Waltham, MA: Ginn), 184–221.

Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. (2005). Three factors in language design. *Linguist. Inq.* 36, 1–22. doi: 10.1162/0024389052993655

Chomsky, N. (2013). Problems of projection. *Lingua* 130, 33–49. doi: 10.1016/j.lingua.2012.12.003

Cinque, G. (2005). Deriving Greenberg's Universal 20 and its exceptions. *Linguist. Inq.* 36, 315–332. doi: 10.1162/0024389054396917

Citko, B. (2014). *Phase Theory: An Introduction*. Cambridge: Cambridge University Press.

Clyne, M. (1987). Constraints on code-switching: how universal are they? *Linguistics* 25, 739–764.

Costa, A., Colomé, À., and Caramazza, A. (2000). Lexical access in Speech production: the bilingual case. *Psicológica* 21, 403–437. Available online at: http://www.uv.es/psicologica/articulos3.00/mono/monoL5.pdf

Croft, W. (2001). *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.

Croft, W. (2002). *Typology and Universals, 2nd Edn*. Cambridge: Cambridge University Press.

Culicover, P., and Jackendoff, R. (2005). *Simpler Syntax*. Oxford: Oxford University Press.

Curcó, C. (2005). Code switching and identity in the discourse of Catalan immigrants in Mexico. *AILA Rev.* 18, 18–40. doi: 10.1075/aila.18.04cur

den Dikken, M. (2006). *Relators and Linkers: The Syntax of Predication, Predicate Inversion, and Copulas*. Cambridge, MA: MIT Press.

DiSciullo, A., Muysken, P., and Singh, R. (1986). Government and code-mixing. *J. Linguist.* 22, 1–24.

Dryer, M. S. (2013a). "Prefixing vs. suffixing in inflectional morphology," in *The World Atlas of Language Structures Online*, eds M. S. Dryer and M. Haspelmath (Leipzig: Max Planck Institute for Evolutionary Anthropology). Available online at: http://wals.info/chapter/26 (Accessed May 23, 2015).

Dryer, M. S. (2013b). "Order of object and verb," in *The World Atlas of Language Structures Online*, eds M. S. Dryer and M. Haspelmath (Leipzig: Max Planck Institute for Evolutionary Anthropology). Available online at: http://wals.info/chapter/26 (Accessed May 23, 2015).

Embick, D., and Noyer, R. (2001). Movement operations after syntax. *Linguist. Inq.* 32, 555–595. doi: 10.1162/002438901753373005

Furukawa, T. (2008). Head/complement relations: portmanteau code-switching between Japanese and English. *Lang. Info. Sci.* 6, 283–292. Available online at: https://www.academia.edu/6084423/Head_complement_relations_Portmanteau_code-switching_between_Japanese_and_English._Language_and_Information_Science_6_283-292

Gafaranga, J. (2000). Medium repair vs. other-language repair: telling the medium of a bilingual conversation. *Int. J. Bilingual.* 4, 327–350. doi: 10.1177/13670069000040030301

González-Vilbazo, K., and López, L. (2011). Some properties of light verbs in code-switching. *Lingua* 121, 832–850. doi: 10.1016/j.lingua.2010.11.011

González-Vilbazo, K., and López, L. (2012). Little *v* and parametric variation. *Nat. Lang. Linguist. Theory* 30, 33–77. doi: 10.1007/s11049-011-9141-5

Green, D. W. (1986). Control, activation, and resource: a framework and a model for the control of speech in bilinguals. *Brain Lang.* 27, 210–223.

Green, D. W., and Li, W. (2014). A control process model of code-switching. *Lang. Cogn. Neurosci.* 29, 499–511. doi: 10.1080/23273798.2014.882515

Grosjean, F. (2008). *Studying Bilinguals*. Oxford: Oxford University Press.

Haegeman, L., and Guéron, J. (1999). *English Grammar*. Oxford: Blackwell.

Haiman, J. (1983). Iconic and economic motivation. *Language* 59, 781–819.

Haiman, J. (1985). *Natural Syntax*. Cambridge: Cambridge University Press

Haspelmath, M. (2007). "Coordination," in *Language Typology and Syntactic Description, Volume II, Complex Constructions, 2nd Edn.,* ed T. Shopin (Cambridge, UK: Cambridge University Press), 1–51.

Haust, D., and Dittmar, N. (1997). "Taxonomic or functional models in the description of code-switching? Evidence from Mandinka and Wolof in African contact situations," in *Codeswitching Worldwide*, ed R. Jacobson (Berlin: Mouton de Gruyter), 79–90.

Hicks, C. (2010). *Morphosyntactic Doubling in Code Switching*. Chapel Hill, NC: University of North Carolina, MA thesis.

Hicks, C. (2012). A dual structure analysis of morphosyntactic doubling in code switching. *Stud. Linguist. Sci.* 2012, 44–57. Available online at: http://hdl.handle.net/2142/35295

Jackendoff, R. (1977). *X-Bar Syntax: A Study of Phrase Structure*. Cambridge, MA: MIT Press.

Jónsson, J. G. (2008). "Preposition reduplication in Icelandic," in *Syntax and Semantics 36: Microvariation in Syntactic Doubling*, eds S. Barbiers, O. Koeneman, M. Lekakou, and M. van der Ham (Bingley: Emerald), 403–417.

Kallulli, D., and Rothmayr, A. (2008). The syntax and semantics of indefinite determiner doubling constructions in varieties of German. *J. Compar. Germanic Linguist.* 11, 95–136. doi: 10.1007/s10828-008-9019-z

Kayne, R. (1994). *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kayne, R. (2004). "Antisymmetry and Japanese," in *Variation and Universals in Biolinguistics*, ed L. Jenkins (Amsterdam: Elsevier B.V.), 3–35.

Kayne, R. (2005). "Some notes on comparative syntax, with special reference to English and French," in *The Oxford Handbook of Comparative Syntax*, eds G. Cinque and R. Kayne (New York, NY: Oxford University Press), 3–69.

Kayne, R. (2009). Antisymmetry and the lexicon. *Lang. Var. Yearb.* 8, 1–31. doi: 10.1075/livy.8.01kay

Kremers, J. (2009). Recursive linearization. *Linguist. Rev.* 28, 135–166. doi: 10.1515/tlir.2009.004

Kremers, J. (2012). The syntax of simultaneity. *Lingua* 122, 979–1003. doi: 10.1016/j.lingua.2012.03.008

Langacker, R. W. (2008). *Cognitive Grammar: A Basic Introduction*. New York, NY: Oxford University Press.

Leung, T. T.-C. (2001). *An Optimality-Theoretic Approach to Cantonese/English Code-Switching*. Unpublished MPhil Dissertation. Hong Kong: The University of Hong Kong.

Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.

Li, D. (2001). L2 lexis in L1: reluctance to translate out of concern for referential meaning. *Multilingua* 20, 1–26. doi: 10.1515/MULTI.2001.001

MacSwan, J. (1999). *A Minimalist Approach to Intra-sentential Code-switching*. New York, NY: Garland.

MacSwan, J. (2000). The architecture of the bilingual language faculty: evidence from intrasentential code switching. *Bilingualism* 3, 37–54. doi: 10.1017/S1366728900000122

MacSwan, J. (ed.). (2014). *Grammatical Theory and Bilingual Codeswitching*. Cambridge, MA: MIT Press.

Mahootian, S. (1993). *A Null Theory of Code-switching*. Doctoral Dissertation, Northwestern University.

Moro, A. (2010). *Dynamic Antisymmetry*. Cambridge, MA: MIT Press.

Muysken, P. (2000). *Bilingual Speech*. Cambridge: Cambridge University Press.

Muysken, P. (2013). Language contact outcomes as the optimization of strategies. *Bilingualism* 16, 709–730. doi: 10.1017/S1366728912000727

Myers-Scotton, C. (1993). *Duelling Languages: Grammatical Structure in Codeswitching*. Oxford: Clarendon Press.

Myers-Scotton, C. (2002). *Contact Linguistics*. Oxford: Oxford University Press.

Myers-Scotton, C. (2006). *Multiple Voices: An Introduction to Bilingualism*. Oxford: Wiley-Blackwell.

Myers-Scotton, C., and Jake, J. L. (2009). "A universal model of code-switching and bilingual language processing and production," in *The Cambridge Handbook of Linguistic Code-Switching*, eds B. E. Bullock and A. J. Toribio (New York, NY: Cambridge University Press), 336–357.

Namba, K. (2012a). *English-Japanese Code-switching and Formulaic Language: A Structural Approach in Bilingual Children's Interactions*. Saarbrücken: LAP Lambert Publishing.

Namba, K. (2012b). Non-insertional code-switching in English-Japanese bilingual children: alternation and congruent lexicalization. *Int. J. Biling. Educ. Bilingual.* 15, 455–473. doi: 10.1080/13670050.2012.665829

Nishimura, M. (1985). *Intrasentential Code-switching in Japanese and English*. Doctoral Dissertation, University of Pennsylvania.

Nishimura, M. (1986). "Intrasentential code-switching: the case of language assignment," in *Language Processing in Bilinguals: Psycholinguistic and Neuropsychological Perspectives*, ed J. Vaid (Hillsdale, NJ: Lawrence Erlbaum), 123–144.

Nishimura, M. (1995). A functional analysis of Japanese/English code-switching. *J. Pragmatics* 23, 157–181.

Pandit, I. (1986). *Hindi-English Code-Switching: Mixed Hindi English*. Delhi: Datta Book Centre.

Park, J. (1990). *Korean/English Intrasentential Code-switching: Matrix Language Assignment and Linguistic Constraints*. Doctoral Dissertation, University of Illinois at Urbana Champaign.

Poplack, S. (1980). Sometimes I'll start a sentence in English y terminol Espanol: towards a typology of code-switching. *Linguistics* 18, 581–618.

Runnqvist, E., Fitzpatrick, I., Strijkers, K., and Costa, A. (2013). "An appraisal of the bilingual language production system: quantitatively or qualitatively different from monolinguals?" in *The Handbook of Bilingualism and Multilingualism*, eds T. K. Bhatia and W. C. Ritchie (Oxford: Blackwell Publishing), 243–265.

Sadock, J. M. (1983). Autolexical syntax: a proposal for the treatment of noun incorporation and similar phenomena. *Nat. Lang. Linguist. Theory* 3, 379–439.

Sadock, J. M. (1991). *Autolexical Syntax: A Theory of Parallel Grammatical Representations*. Chicago, IL: Chicago University Press.

Sankoff, D., and Poplack, S. (1981). A formal grammar for code-switching. *Papers Linguist.* 14, 33–46.

Sankoff, D., Poplack, S., and Vanniarajan, S. (1990). The case of the nonce loan in Tamil. *Lang. Var. Change* 2, 71–101.

Santorini, B., and Mahootian, S. (1995). Codeswtiching and the syntactic status of abnominal adjectives. *Lingua* 96, 1–27.

Shim, J. Y. (2013). *Deriving Word Order in Code-Switching: Feature Inheritance and Light Verbs*. Doctoral Dissertation, The City University of New York.

Slabbert, S., and Myers-Scotton, C. (1996). The structure of Tsotsitaal and Iscamtho: code-switching and in-group identity in South Africaan townships. *Linguistics* 34, 317–342.

Stolz, T. (1996). Grammatical hispanisms in Amerindian and Austronesian languages: the other kind of transpacific isoglosses. *AMERINDIA* 21, 1–22.

Takagi, M. (2007). Alternational code-switching in the story-telling of English-Japanese bilingual children narratives. *J. Inq. Res.* 86, 1–18. Available online at: https://opac1.kansaigaidai.ac.jp/webopac/bdyview.do?bodyid=DB00000198

Taylor, J. (2002). *Cognitive Grammar*. Oxford, UK: Oxford University Press.

Toribio, A. J. (2001). On the emergence of bilingual code-switching competence. *Bilingualism* 4, 203–231. doi: 10.1017/s1366728901000414

Wilson, D., and Carston, R. (2007). "A unitary approach to lexical pragmatics: relevance, inference and *ad hoc* concepts," in *Pragmatics*, ed N. Burton-Roberts (Basingstoke; New York, NY: Palgrave Macmillan), 230–259.

Woolford, E. (1983). Bilingual code-switching and syntactic theory. *Linguist. Inq.* 14, 520–536.

Wu, Y. J., and Thierry, G. (2010). Investigating bilingual processing: the neglected role of language processing contexts. *Front. Psychol.* 1:178. doi: 10.3389/fpsyg.2010.00178

Grammatical Encoding in Bilingual Language Production: A Focus on Code-Switching

Mehdi Purmohammad*

Center for the Study of Language and Society (CSLS), University of Bern, Bern, Switzerland

In this study, I report three experiments that examined whether words from one language of bilinguals can use the syntactic features form the other language, and how such syntactic co-activation might influence syntactic processing. In other words, I examined whether there are any cases in which an inherent syntactic feature a lexical item is inhibited and the syntactic feature that belongs to the other language is used, instead. In the non-switch condition in Experiments 1 and 2, Persian-English bilinguals described pictures using an adjective–noun string from the same language requested. In the switch condition, they used a noun and an adjective from the other language. In the switch condition in Experiment 3, participants used only the adjective of a noun phrase from the other language. The results showed that bilinguals may inhibit the activation of a word's syntactic feature and use the syntactic property from the other language, instead [e.g., pirāhane (shirt-N) black]. As the combinatorial node (the node that specifies different kinds of syntactic structures in which a word can be used) of a used adjective retains activation at least temporarily, bilinguals are more likely to use the same combinatorial node even with an adjective from the other language. Cross-language syntactic interference increased in the switch conditions. Moreover, more inappropriate responses were observed when switching from bilinguals' L2 to L1. The results also revealed that different experimental contexts may lead to different patterns of the control mechanism. The results will be interpreted in terms of Hartsuiker and Pickering's (2008) model of syntactic representation.

Keywords: bilingualism, bilingual language production, code-switching, grammatical encoding, syntactic processing

INTRODUCTION

Code-switching (CS) is defined as a change from one language of a bilingual speaker to another in the same utterance or conversation (Hamers and Blanc, 1989). CS is a common language phenomenon that occurs in bilinguals' speech production. Example (1) shows CS between English and Spanish:

(1) Dónde está ese paño blue?

'Where is that blue cloth?'   Arias and Lakshmanan (2005, p. 104)

The CS phenomenon has been widely discussed in a variety of fields. In comparison with all other contact phenomena of interest, CS "has arguably dominated the field" (Bullock and Toribio, 2009, p. 1). Psycholinguistic research on aspects of bilingual language production has

focused on general modeling issues (e.g., de Bot, 1992; de Bot and Schreuder, 1993), the control of processing (e.g., Green, 1993, 1998), and the formulation of output (e.g., Myers-Scotton, 1993, 2002) (Karousou-Fokas and Garman, 2001). In all approaches, the CS data are viewed as important sources of evidence. Studies on CS can help psycholinguists, for instance, find whether one of the two languages is deactivated while the other language is being activated, and "how incoming signals are channeled to their appropriate decoding system for interpretation (e.g., input switch)" (Paradis, 1993, p. 135).

Code-switching in constructions containing an adjective has received a lot of attention in structural linguistics. Most structural approaches to CS look for formulating some constraints on CS. For about three decades now, the main aim of positing the constraints has been to formulate the interaction between the two grammars of a bilingual speaker in CS (Mahootian, 2006). Some earlier research (e.g., Pfaff, 1979; Sankoff and Poplack, 1981) proposed that CS is not allowed at points where the two languages in contact do not share the same word order representation (see MacSwan, 2009). Accordingly, "adjective/noun mixes must match the surface word order of both the language of the adjective and the language of the head noun" (Pfaff, 1979, p. 306). In this view, since Persian and English do not share the same adjective–noun order, switching inside NPs is prohibited. Some researchers (e.g., Aguirre, 1976; McClure, 1977, 1981) argued that switching inside NPs is possible so long as the placement rule of the adjective language is met. The Equivalence Constraint Model of Code-switching (Poplack, 1980; Sankoff and Poplack, 1981) stipulates that language switching tends to occur at points where the two languages have the same word order representation. Thus, according to Poplack (1980) and Sankoff and Poplack (1981), since the syntactic rule of one of the two languages is violated in the Persian-English switches inside the NP structures, switches do not occur.

Purmohammad (2015) investigated the grammatical encoding in code-switched utterances. He collected 2293 min of a popular TV show. Persian-English bilinguals freely inserted English words into their Persian utterances. 962 code-switched utterances were found. He reports that 210 switched words were adjectives. In 10% of the cases, Persian-English bilinguals used English adjectives after the Persian nouns.

Cantone and MacSwan (2009) investigated how linguistic properties relevant to determining surface word order for adjectival constructions are resolved in CS contexts in which languages with different word order are involved. In line with this, 10 participants gave their grammaticality judgments for the mixed utterances involving determiners, adjectives, and nouns (e.g., in un Bett nuovo meaning a bed new) by determining whether each utterance was well-formed or not. The results of the study showed that whenever the language of the adjective was reflected in the word order of the mixed utterances, participants judged them to be acceptable; whereas those mixed utterances in which the language of adjective was not matched were judged to be ill-formed. Some researchers (e.g., Pandit, 1990) assumed that the language of the head noun determines the syntactic properties of its complements; however, Nartey (1982,

cited in Cantone and MacSwan, 2009) assumed that in the Adanme-English CS, the Adanme determiner determines noun-adjective order. Belazi et al. (1994) claim that the language of the adjective determines the adjective–noun order. As we will see later, the results of the present study are inconsistent with the constraints proposed on adjective–noun switches; however, we will not go into more details here (see Gil et al., 2012 for more discussion).

Bilingual speakers know two different languages and hence they know two different grammatical systems. For example, one of the two languages of a Persian-English bilingual speaker uses post-nominal adjectives (adjectives follow nouns) whereas the other language (English) uses prenominal adjectives (adjectives precede nouns). Although ample evidence has led researchers to assume that the two languages are co-activated during lexical processing, a fundamental question is whether the parallel activation of the two languages leads to interference (Hatzidaki et al., 2011). One group of researchers assumes that although the two languages of bilinguals are activated during sentence production, the non-target language does not affect the target language. For example, La Heij (2005) proposed that the intended language acts as a language cue. It ensures that lexical items in the intended language reach a higher activation level than their equivalent translations in the non-intended language. The second group of researchers suggests that activation of the non-intended language can influence lexical processing in the target-language (see Costa, 2005 for review). For example, Costa et al. (2006) tested for the lexical bias effect (LBE). This effect shows "feedback between the phonological and lexical levels of representation during speech production" (p. 972). The LBEs suggest that feedback existing in second-language production extends across the two languages of a bilingual speaker. They conclude that representations of both languages are recruited in bilingual language processing even when only one language is used.

As stated above, there is compelling evidence (e.g., Francis, 2005; Kroll et al., 2006, 2008; Voga and Grainger, 2007) indicating that aspects of the two languages of bilinguals are activated during both unilingual and bilingual modes (see Grosjean, 2008 for language mode account). Thus, we expect syntactic interference from the non-target language. Although the results of studies has provided the researchers the evidence to assume that components of the two languages (e.g., syntax, phonology) are activated during language processing, it remains contentious what exactly means by interference, for instance, the syntactic interference, especially from a processing perspective. More importantly, it remains unclear how the processor operates during language interference. For example, what language processing mechanism underlies the sentence in which a bilingual uses a prenominal adjective (e.g., Spanish "chiquita" meaning small) post-nominally? (see example 2).

(2) I went to the house CHIQUITA.

　　I went to the little house.　　　　　(Pfaff, 1979, p. 307)

This study examines whether words from language A can use the syntactic features form language B and how such syntactic co-activation might influence syntactic processing. To put it

differently, the main aim is to examine whether there are any cases in which an inherent syntactic feature (e.g., post-nominality) of a lexical item (e.g., an adjective) is inhibited and the syntactic feature that belongs to the other language is used, instead. If this were the case, how such linguistic behavior could be captured within a model of bilingual language production.

The present study reports three experiments that investigate the processing of adjective–noun strings in code-switched utterances. More specifically, I examine how the activation of adjective placement rule from the non-target language may affect the syntactic processing of the structures containing a noun and an adjective. In all three experiments, participants use adjective–noun strings in order to name pictures. If their language productions differ with respect to using syntactic features in three experiments, I will discuss what factors might cause such differences. If the grammatical features of the non-target language affect the target language, for instance, using an English adjective post-nominally (e.g., "ketābe different" lit. "book different"), this would give evidence to suggest that bilingual's two language systems interfere during language processing. Finally, I examine how the results of the present study might be integrated with Hartsuiker and Pickering's (2008) integrated model of syntactic representation.

According to Hartsuiker and Pickering's (2008) model, bilingual speakers have an integrated lemma stratum. It is assumed that lemmas – the base form of each word- from the two languages are represented in an integrated network. Each lemma node (e.g., red in English or *qermz* in Persian) is linked to one conceptual node [RED(X,Y)] at the conceptual stratum, to one category node (e.g., adjective, noun), to combinatorial nodes (e.g., prenominal or post-nominal adjective), and to one language node (e.g., English, Persian) in their integrated network. In this model, category nodes specify grammatical categories (e.g., adjective) and combinatorial nodes specify different kinds of syntactic structures in which a word can be used (Bernolet et al., 2007). One of the important aspects of the model is that featural, combinatorial, and category nodes are shared in a way that reduce redundancy (Cleland and Pickering, 2003). Accordingly, the lemma nodes such as "nice" and "brown" are both linked to the same category node (adjective) and combinatorial node (prenominal).

Cross-linguistic grammatical effects and lexical switching are predicted in this model, because in this model both meaning and syntax of lexical items are points of contact across languages (Hartsuiker et al., 2004). Thus, according to the model's prediction it is possible that a Persian-English bilingual speaker selects a Persian construction (e.g., a noun-adjective word order string) when using an English adjective (e.g., a book red). However, no effect of language proficiency on cross-linguistic influences was predicted by the model (Hartsuiker and Pickering, 2008).

Given that Persian uses adjectives post-nominally while English generally uses adjectives prenominally, it seems that adjective placement is suitable for the purpose of the study, because the results may better show how the syntactic components of the two languages of bilinguals interact during speech production compared to the situation in which both languages use the same adjective placement rules. In this study, adjectives are used either prenominally or post-nominally in their corresponding languages. When interference occurs, an adjective is likely to cede its combinatorial feature (prenominal or post-nominal) to the other combinatorial feature. A model of bilingual syntactic representation needs to explain how the production of a lexical element is influenced by the syntactic properties of the other language.

Investigating syntactic interference is crucial because this phenomenon permits us to know how the grammars of the two languages are represented in bilinguals' memory; how the grammars of the two languages interact during production; how grammatical functions are assigned to concepts, and more importantly how the mental lexicon and syntactic encoding interface in bilingualism (Hartsuiker and Pickering, 2008). All three experiments reported in this study include switching tasks. Since the aim of the study is to test whether there are any cases in which bilingual speakers use the grammar of one language and the words from the other language, it seems that language switching tasks are suitable for the purpose of the study, because when a bilingual speaker switches between the two languages, he or she has to consider using two different grammatical systems in a single utterance.

EXPERIMENTS

The present study consists of three experiments. In all experiments a picture-naming task was used. In each trial, participants were presented with a sentence fragment along with a picture depicted above the sentence fragment. In Experiment 1, in the non-switch conditions, participants described pictures using an adjective–noun string from the language of the sentence fragment. In the switch conditions, however, they completed the sentence fragments using a noun and an adjective from the other language (see **Table 2**, for sample items used in Experiment 1). In Experiment 2 in the switch conditions, participants were presented with a sentence fragment in language A along with a picture depicted above it. A noun phrase including q noun and an adjective from language B was printed above the target picture as well (see **Table 5**, for sample items used in Experiment 2 and Appendix A for the items used in Experiment 2). Participants had to use the translation-equivalents of the noun phrase in order to describe pictures. In the non-switch conditions, however, they used both nouns and adjectives from the language of the sentence fragment. In each trial in Experiment 3, participants were presented with a sentence fragment along with a picture depicted above it. In the switch conditions, participants used only the adjectives of noun phrases from the other language; however, they used the noun from the language of the sentence fragment. In the non-switch conditions, they had to use a noun and an adjective from the language of the sentence fragment (see Appendix B for the items used in Experiment 3).

All experiments consisted of two main conditions (the switch and non-switch conditions) and four different sets of items: the Persian set, the Persian-English set, the English set and the English-Persian set. The Persian and English sets of

items represented the non-switch conditions and the Persian-English and English-Persian sets of items represented the switch conditions. The experiments, thus, had a 2x2 experimental design for language task (the switch vs. non-switch condition) and language (Persian vs. English; Persian-English vs. English-Persian).

Consistent with Hatzidaki et al. (2011), it is hypothesized that since the two languages of bilingual speakers are activated during language production, the grammatical system of the non-target language may affect the production of the target language. Moreover, it is hypothesized that more inappropriate responses in which a word from language A uses a syntactic feature from language B (e.g., "marde tall" lit. "man tall") are made in the tasks that involve switching (i.e., in bilingual contexts) than in the unilingual contexts involving no switching, because in the switch conditions the two languages of a bilingual speaker must inevitably be activated and that in the switch conditions both languages are activated to a greater degree compared to the non-switch conditions.

EXPERIMENT 1: SENTENCE COMPLETION TASK 1

Method
Participants
Thirty six Persian (L1)-English (L2) bilinguals took part in the experiment. Participants were recruited through advertisements which clearly stated proficiency in both Persian and English as prerequisite. They were paid six pounds for their participation. Eighteen of them were Ph.D. students at Heriot-Watt University or the University of Edinburgh. Eleven participants hold master's degrees from the UK universities. Two of them were university professors. Five participants were high school students in Edinburgh. They all reported having normal vision. Their self-ratings of English language skills (speaking and listening) and the results of the English proficiency test demonstrated that the participants were fluent in English. The median age of the participants was 30.5 years with a median length of residence of 8 years in UK. **Table 1** shows the participants' background characteristics in all three experiments reported in this study.

Materials
Thirty-two sentence fragments were created. The 32 sentence fragments included eight items from the Persian set, the English set, the Persian-English set, and the English-Persian set. In each trial, the name of a common object was omitted. Thirty-two unique pictures were presented in the place of the omitted objects. For the Persian set, the green outlined pictures were used to satisfy Persian as the response language. Then a mixture of eight green outlined pictures with eight Persian sentence fragments was used for the Persian set. For the English set, the orange outlined pictures were used to satisfy English as the response language. A mixture of eight orange outlined pictures with eight English sentence fragments was used for the English set. The English-Persian set was created by combining the English sentence fragments with the green outlined pictures.

The Persian-English set was created by combining the Persian sentence fragments with the orange outlined pictures. In each experiment, 32 highly frequent nouns (16 nouns for the English set and 16 nouns for the Persian set) and 32 highly frequent adjectives (16 adjectives for the English set and 16 adjectives for the Persian set) were used. It is common to use background-color-cueing procedure in language switching studies (see Meuter and Allport, 1999; Costa and Santesteban, 2004; Kootstra et al., 2010; Broersma, 2011).

Two randomized versions of the same presentation list were constructed. Each list included 32 items. Sixteen Persian sentence fragments were constructed and their English translations were used for the English set. A group of five Persian-English speakers was asked to check for the accuracy of English sentences. Pictures were identical in all sets. Sixteen Persian-English sentence fragments were provided and their English translations were used for the English-Persian set. Each list contained eight items from each set (the Persian set, the English set, the Persian-English set, and the English-Persian set). Then Experiment 1 included 16 switch conditions and 16 non-switch conditions. **Table 2** shows sample items used in Experiment 1.

Since the English sentence fragments were the translations of Persian sentence fragments, each list was designed so that participants did not receive two semantically identical items. Trials were in randomized order.

There is a concern that different classes of adjectives may work differently (Sobin, 1984). In the present study, different types of adjectives (e.g., color, feeling, appearance, shape, size) were used; however, Sobin (1984) used color adjectives only.

Procedure
Before doing the experiments, participants were asked some demographic questions including name, age, sex, and the number of the years they used English in their daily life. Prior to the experiments, participants were given four practice trials in order to familiarize themselves with the experimental tasks. Instructions were given in Persian. Participants were informed that their speech would be recorded. Each participant was tested individually. They sat in front of the same laptop and completed the sentence fragments.

TABLE 1 | Participants' characteristics in Experiments 1–3.

Measures	EXP 1 N = 36	EXP 2 N = 37	EXP 3 N = 29
Age	30.5	29.72	29.20
Self-rated speaking ability in English (seven-point scale)	5.55	5.59	5.55
Self-rated listening ability in English (seven-point scale)	5.69	5.72	5.62
English language proficiency test mark (the highest score: 25)	21.75	21.86	22.06
Years of English language use in daily life	8.62	8.87	8.77
Self-reported amount of code-switching (five-point scale)	2.52	2.53	2.61

EXP: Experiment, N: number of participants.

TABLE 2 | Sample items used in Experiment 1.

Sets of items	Sample items
The Persian set	مینا رو برام آورد.*
The Persian-English set	مینا رو برام آورد.
The English set	Judy carried the for me
The English-Persian set	Judy carried the for me

*The table shows the basic design used in Experiment 1. Two semantically identical items were not used in a single list. *Mina carried the "heavy bag" for me.*

In each trial, a sentence fragment along with a picture depicted above it was presented to the participants. Participants were instructed to read the entire sentence fragment out loud and to fill in the missing part. In order to describe the pictures presented in the place of the omitted objects (see **Table 2** for sample items used in Experiment 1), they had to use a noun and an adjective. By doing so, they completed 32 sentence fragments. While the green outlined pictures showed that Persian should be the response language, the orange outlined pictures showed that English should be the response language. Therefore, in the non-switch conditions, if the sentence fragments were in Persian and the pictures had a green background color, participants had to use a Persian noun and an adjective to complete the sentence fragments. In the switch conditions, when the sentence fragments were in Persian and the pictures had an orange background color, they had to use an English noun and an adjective to complete them. In the same way, if the sentence fragments were in English and the pictures had a green background color, participants had to use a Persian noun and an adjective to complete the sentence fragments. They were told that there was no preferable way of doing the task.

A 25-items cloze test was constructed to rate participants' English language proficiency. Participants were instructed to fill in the blanks with the most appropriate English words.

Scoring and Data Analysis

Three different categories were used to score participants' responses. Responses were scored as "appropriate" when participants completed the sentence fragments as requested (i.e., using an English adjective prenominally and using a Persian adjective post-nominally). Responses were scored as "inappropriate" when they did not complete the sentence fragments as requested. Then a response that used an English adjective post-nominally (e.g., "chiz-e different" lit. "thing different") is considered as an "inappropriate" response. Responses were scored as "other" for all other completions. For example, if participants failed to complete a sentence fragment, it would be scored as "other." Moreover, all responses had to use a noun-adjective string only. All other strings (e.g., a lot of books) were scored as "other" and omitted from the analyses.

The scoring criteria need more clarification. Responses were scored as "appropriate" when participants used the correct adjective placement rule of the language that the adjective belongs to (e.g., "tall mard" lit. "tall man"). Accordingly, prenominality is considered as an inherent feature of adjectives in English and post-nominality is considered as an inherent feature of adjectives in Persian. Then a response that used an English adjective post-nominally (e.g., "chiz-e different" lit. "thing different") is considered as "inappropriate." Note that I did not consider the structural accounts on the syntactic structure of CS involving adjectival constructions, because all the responses in the switch conditions are inconsistent with the structural accounts in which language switching is not allowed at points where the two languages do not share the same word order representation (see Gil et al., 2012). Moreover, there is ample evidence indicating that neither the head noun nor the adjective, nor the language of the determiner (e.g., a, the) determine the adjective–noun order (see Introduction; Cantone and MacSwan, 2009 for review). I am concerned whether or not adjectives use the adjective placement rule of the language they belong to. Thus, "appropriateness" here does not mean that participants used the correct adjective–noun order in language A or B, because when the two languages of bilinguals use different adjective–noun ordering, we always expect that the switch containing a noun and an adjective does not respect the language-specific requirement of one of the two languages involved.

Similar to Hatzidaki et al. (2011) and Selles (2011), a linear mixed effect was used to test whether the inappropriate responses were affected by language task (the switch and non-switch conditions), language proficiency, source language, target language, and participants' self-ratings of their speaking and listening skills. Using appropriate and inappropriate responses as the dependent variables and experimental items and participants as random effects, first a null model was created. To find the model with the best fit, predictors were added to the model individually. Then using χ^2-tests, the models were compared to see whether adding the predictors contributed significantly to the model.

Results

Overall, 1152 sentence fragments including 576 switched and 576 non-switch utterances were completed by the participants. There were 10 (0.86%) "other" responses and removed from the analyses. The analysis is based on the remaining 1142 sentence fragment completions. The results of Experiment 1 showed that appropriate responses occurred more frequently (98%) than inappropriate responses (2%). The number of appropriate

responses was almost the same in the non-switch conditions (98%) and in the switch conditions (97%). Moreover, the results demonstrated that inappropriate responses occurred more frequently in the switches from L2 to L1 (78.52%) than from L1 to L2 (21.42%). **Table 3** reports the participants' responses per condition.

Using a linear mixed effect model, a baseline model was created using participants and items as random effects. The logistic variant was used. Items and participants were used as random slopes. I incrementally added predictors to the base line model and χ^2-tests were conducted to determine which of the predictors attributed to the model of best fit (see **Table 4**). Language task, target language, and source language were tested individually as predictors. Language task and target language were individually significant but source language was not significant. Finally, both language task and target language were added to the base model as predictors and the results were highly significant. χ^2-tests showed that the model of best fit used language task and target language as predictors.

As the language task variable is a combination of the two other variables (source language and target language), it may be redundant to include it as a predictor. Thus, it would be sufficient to consider only source language and target language as predictors. Dropping language task from the data analysis yields the following results: no significance in target language × source language interaction, and a main effect of target language ($p < 0.003$). The results also indicated that the language × condition (the switch/non-switch condition) interaction was not significant ($p > 0.7$).

To test to see whether language proficiency put an effect on responses, further predictors were added based on the rating of participants' proficiency levels. English proficiency tested in interaction with experimental predictors yielded the following results: no significance of self-rated language proficiency × target language, language proficiency × source language, self-rated speaking proficiency × target language, self-rated speaking proficiency × source language, self-rated listening proficiency × target language, or self-rated listening proficiency × source language (see **Table 1** for different measures of language proficiency level).

TABLE 3 | Experiment 1: participants' responses in the switch and non-switch tasks.

Language task	Responses			
	Sum	Appropriate	% inappropriate	Omission
Non-switch tasks	576	570	4 (0.70)	2
Persian	288	283	4 (100)	1
English	288	287	0 (0.00)	1
Switch tasks	576	554	14 (2.43)	8
Persian-English	288	280	3 (21.42)	5
English-Persian	288	274	11 (78.57)	3

LT, Language task; Omission: responses scored as other, % inappropriate: the percentage of inappropriate responses (responses scored as other were not included).

Discussion

In sum, the results of Experiment 1 demonstrated that the adjective–noun order of the intended language was a strong predictor of participants' performance both in the switch and non-switch conditions. However, the results revealed that as both languages of bilingual speakers were co-activated, participants showed interference from the non-target language on the target language. Responses were not affected by participants' levels of language proficiency.

EXPERIMENT 2: SENTENCE COMPLETION TASK 2

To get a better picture of the nature of the syntactic processing in code-switched utterances, an additional sentence completion task (Experiment 2) was designed. Experiment 2 investigates whether similar results would occur with a different task in which participants use the translation equivalents of the noun phrases printed above pictures in the switch conditions.

Method
Participants
Thirty-seven participants took part in Experiment 2. Thirty-six of them were from the same population as Experiment 1. **Table 1** demonstrates the participants' characteristics.

Materials
Thirty-two sentence fragments were created. Thirty-two unique pictures were presented in the place of the omitted objects (see Appendix A for the items used in the experiment). The pictures were identical across the four language sets. The main difference between the switch and non-switch trials was that in the switch conditions noun phrases from the base language (the language of the sentence fragments) were printed above the target pictures. Participants had to use the translation-equivalents of the noun phrases printed above the target pictures. But in the non-switch trials, they had to use a noun and an adjective from the base language to describe pictures. As in Experiment 1, the 32 sentence fragments included eight items from the Persian set, the English set, the Persian-English set, and the English-Persian set. Then Experiment 2 consisted of 16 switch trials and 16 non-switch trials. **Table 5** shows sample items used in Experiment 2.

Two randomized versions of the same presentation list were constructed. Since the English sentence fragments were the translations of the Persian sentence fragments, the lists were arranged so that not each participant received two semantically identical items.

Procedure
Participants were instructed that in the switch conditions they would first read the noun phrases printed above the target pictures. To describe pictures, they had to use the translation-equivalents of the noun phrases printed above the target pictures. Participants were told that in the non-switch trials, they had to use a noun and an adjective from the language of the sentence fragment (base language). Prior to the experiments, participants

TABLE 4 | Models of responses in Experiment 1.

Predictor	Estimate	SE	z-value	p
Language task as main predictor: $\chi^2(1) = 5.618$, $p = 0.018$, $N = 1142$				
(Intercept)	−10.445	1.727	−6.049	< 0.001
Language task	1.479	0.685	2.159	0.031
Target language as main predictor: $\chi^2(1) = 8.578$, $p = 0.003$, $N = 1142$				
(Intercept)	−5.646	1.536	−3.675	< 0.001
Target language	−1.846	0.723	−2.552	0.011
Language task and target language as predictors: $\chi^2(1) = 15.601$, $p < 0.001$, $N = 1142$				
(Intercept)	−8.284	1.938	−4.275	< 0.001
Language task	1.508	0.642	2.349	0.019
Target language	−1.899	0.706	−2.689	0.007

TABLE 5 | Sample items used in Experiment 2.

Sets of items	Sample items
The Persian set	خرید.* توماس برای تولد سارا
	گردن بند گرون
The Persian-English set	خرید. توماس برای تولد سارا
The English set	Thomas bought Sarah an ⬜ for her birthday
	expensive necklace
The English-Persian set	Thomas bought Sarah an ⬜ for her birthday

*The Table shows the basic design used in Experiment 2. Two semantically identical items were not used in a single list. *Thomas bought Sarah an "expensive necklace" for her birthday.*

were given four practice trials in order to familiarize themselves with the experimental tasks. Participants were informed that their speech would be recorded. They were instructed entirely in Persian.

Scoring and Data Analysis

The scoring and data analysis were identical to those described in Experiment 1.

Results

Overall, 1185 sentence fragments consisting of 592 switched and 592 non-switched utterances were completed by the participants. There were 10 (0.84%) "other" responses and discarded from the

analysis. The following analysis is based on the remaining 1175 responses. The results showed that the global pattern of responses was identical to those in Experiment 1. Similar to Experiment 1, in most cases (98%) participants used the correct adjective placement rules. Participants produced more appropriate responses (99%) in the non-switch conditions than in the switch conditions (96%). **Table 6** shows participants' responses per condition. The results also revealed that inappropriate responses occurred more frequently in switches from L2 to L1 (89.47%) than in switches from L1 to L2 (10.52%).

The results were calculated as described in Experiment 1. Target language and source language were individually significant, but language task (trial type) was not. When target language and source language were both added as predictors they had significant effects on model. As in Experiment 1, χ^2-tests were conducted to determine the model of best fit (see **Table 7**). With the χ^2-tests, it was found that the model with source language and target language as predictors was the model of best fit. Dropping language task from the data analysis did not change the results. There was a main effect of target language and source language ($p < 0.005$). Having removed language task as a predictor, target language is the model of best fit. The results revealed that the language × condition interaction was not significant ($p > 0.1$). Language proficiency was tested in interaction with the experimental predictors. Similar to Experiment 1, neither language proficiency, nor self-rating of

TABLE 6 | Experiment 2: participants' responses in the switch and non-switch tasks.

Language task	Responses			
	Sum	Appropriate	% inappropriate	Omission
Non-switch tasks	592	582	1 (0.16)	9
Persian	296	292	0 (0.00)	4
English	296	290	1 (100)	5
Switch tasks	592	572	19 (3.20)	1
Persian-English	296	293	2 (10.52)	1
English-Persian	296	279	17 (89.47)	0

LT, Language task; Omission: responses scored as other, % inappropriate: the percentage of inappropriate responses (responses scored as other were not included).

TABLE 7 | Models of responses in Experiment 2.

Predictor	Estimate	SE	z-value	p
Target language as main predictor: $\chi^2(1) = 7.995$, $p = 0.005$, $N = 1174$				
(Intercept)	−5.106	1.807	−2.826	0.005
Target language	−2.680	1.148	−2.336	0.020
Source language as main predictor: $\chi^2(1) = 7.997$, $p = 0.005$, $N = 1174$				
(Intercept)	−13.145	2.431	−5.406	< 0.001
Source language	2.680	1.147	2.336	0.020
Source language and target language as predictors: $\chi^2(1) = 19.318$, $p < 0.001$, $N = 1174$				
(Intercept)	−9.009	2.127	−4.237	< 0.001
Source language	2.568	0.838	3.064	0.002
Target language	−2.567	0.838	−3.064	0.002

speaking skill, nor self-rating of listening skill improved the model.

Discussion

The purpose of Experiment 2 was to examine whether a different language task (a translation task) would yield different responses. In the switch trials participants used the translation-equivalents of the noun phrases printed above pictures in order to describe the target pictures. In the non-switch trials; however, they used a noun and an adjective from the language of the sentence fragments. The results showed that as in Experiment 1, in most cases participants used the Persian adjectives post-nominally and the English adjectives prenominally. The results, however, revealed that the intrinsic syntactic feature (the prenominal or post-nominal features) of an adjective can be inhibited and the syntactic feature from the other language can be used, instead. Inappropriate responses were not affected by participants' levels of language proficiency.

Experiment 2 is important because in this experiment, again participants had to use a noun and an adjective from the base language (the non-switch condition) or from the other language (the switch condition). What the results may suggest above and beyond Experiment 1 is that when both the noun and adjective are from the same language, adjectives were appropriately located. The results show that the context or the task in which a word is produced may affect the syntactic processing during sentence production.

EXPERIMENT 3: SENTENCE COMPLETION TASK 3

Experiment 3 examines whether using the syntactic features (combinatorial nodes) from the other language enhances when only adjectives from the other language have to be used in the switch conditions. Experiment 3 used the same design as Experiment 1 except that in the switch conditions participants used only adjectives from the other language.

Method
Participants

Twenty-nine subjects from the same population as Experiment 1 were recruited to participate in this study (see **Table 1** for

participants' characteristics). They were tested 2 weeks after they had participated in Experiments 1 and 2.

Materials and Designs

These were identical to those described in Experiment 1 except that 10 sentence fragments were replaced by new sentence fragments (see Appendix B for materials used in this experiment). Such replacement was done so that participants would feel that they were performing an experiment that used a different task and different materials from Experiment 1.

Procedure

Participants were seated in front of a laptop and completed the sentence fragments. Experiment 3 used the same background color cues as in Experiment 1 (see the procedure described in Experiment 1). As in Experiment 1, participants were instructed to use a noun–adjective string to describe the target pictures. The main difference between Experiment 1 and Experiment 3 was that in Experiment 3 participants were told to use only adjectives of noun phrases from the other language in the switch trials. In the non-switch trials they had to describe pictures using both adjectives and nouns from the same language depending on which language was requested. Prior to the experiment, participants were given eight practice trials in order to familiarize themselves with the experimental task. Instructions were given in Persian. Participants were informed that their speech would be recorded.

Scoring and Data Analysis

The scoring and data analysis were identical to those described in Experiment 1.

Results

Overall, 928 sentence fragments consisting of 464 switched and 464 non-switched sentence fragments were completed by the participants. Twenty-eight (3%) of the responses were scored as "other" and removed from the analysis. Then the analysis is based on the remaining 900 sentence fragment completions. In sharp contrast to Experiments 1 and 2 in which the grammar of the other language did not considerably affect participants' responses, the syntactic feature of the other language significantly affected participants' responses. The results showed that participants used the adjective placement rule from the other language in (28%) of

the responses. Inappropriate responses occurred more frequently in the switch conditions (93%) than in the non-switch conditions (7%). **Table 8** shows participants' responses per condition. The results demonstrated that in the switch conditions inappropriate responses occurred more frequently in switches from L2 to L1 (65.27%) than in switches from L1 to L2 (36.82%).

As in Experiments 1 and 2, χ^2-tests were conducted to determine the model of best fit (see **Table 9**). The results indicated that language task (trial type) was highly significant ($p < 0.001$). Adding both language task and target language as predictors improved the model significantly. Then target language affects the responses when language task is taken into account. When language task is removed from the data analysis, there was a significant interaction between target language and source language ($p < 0.001$) in Experiment 3.

As in Experiments 1 and 2, language proficiency was tested in interaction with experimental predictors. No significance of English proficiency $\times$ source language, proficiency $\times$ target language, self-rated speaking proficiency $\times$ source language, and self-rated listening proficiency $\times$ source language interaction was observed in Experiment 3. The results revealed that the language $\times$ condition interaction was significant ($p > 0.4$). The rated self-rated speaking proficiency $\times$ target language interaction was significant ($p < .002$). Target language $\times$ self-rated speaking proficiency is model of best fit.

The results clearly indicate that participants may inhibit the syntactic properties of one language and use the syntactic feature from the other language.

Discussion

When participants were asked to describe pictures using both a noun and an adjective from the same language or from the other language in the switch and non-switch trials respectively (see Experiments 1 and 2), they used the correct adjective placement feature of the intended languages in most cases. But in Experiment 3, when they were asked to use only the adjectives of the NP structures from the other language, participants were considerably blind to their uses of the combinatorial nodes (adjective placement rule), suggesting that in Experiment 3, adjectives had much less syntactic restrictions to find their positions in noun phrase structures compared to Experiments 1

and 2. That is, participants' choices of the combinatorial nodes of adjectives were more volatile in Experiment 3 compared to Experiments 1 and 2. Using syntactic features from the non-target language was stronger under some linguistic contexts than the others. While language task had a significant effect on participants' responses, with the exception of the interaction of self-rated speaking proficiency $\times$ target language, no significant effect of language proficiency on cross-linguistic influence was observed.

GENERAL DISCUSSION

The main aim of the present study was to examine whether there are any cases in which an inherent syntactic feature of a lexical item is inhibited, and the syntactic feature that belongs to the other language is used, instead. It was hypothesized that since the two languages of bilingual speakers are co-activated during language production, the grammatical system of the non-target language may affect the production of the target language (see Schwartz and Kroll, 2006). The results, especially from Experiment 3, confirm the main hypothesis of the study. The results showed interference with respect to combinatorial processing. However, cross-linguistic influences affected differentially by whether only adjectives were switched or both the nouns and adjectives of noun phrases were switched. In Experiments 1 and 2, adjectives sometimes used the syntactic feature (i.e., the combinatorial node) from the other language. In Experiment 3, however, participants used the adjective placement rule from the other language more frequently.

It was also hypothesized that more inappropriate responses are made in the switch tasks than in the non-switch tasks. The results indicated that in all experiments, inappropriate responses occurred more frequently in the switch conditions than in the non-switch conditions.

The results of the experiments, especially Experiment 3, demonstrated interference between bilinguals' two language systems during speech production. The results indicated that both languages are co-activated in bilingual language production and that bilingual speakers may use the grammar of one language and the word from the other language. The results are consistent with Nicoladis' (2006) study. She examined whether overlap/ambiguity of adjective–noun strings in English and French leads to transfer. In her study, French-English preschool bilingual children named pictures using an adjective–noun string. Their responses were compared to English and French monolingual children. The results of the study showed that bilinguals made more reversals of pre-nominal French adjectives (e.g., "une personne grand" lit. "a person big") than monolingual peers. Moreover, they reversed more post-nominal adjectives (e.g., "un ray'e dinosaure" lit. "a striped dinosaur") than monolingual children. However, more adjective reversal occurred in French, because French uses two adjective–noun orders. The researcher views cross-linguistic transfer as "an epiphenomenon of speech production" (p. 26).

In all three experiments participants used an adjective–noun string to describe pictures; however, their language production

TABLE 8 | Experiment 3: participants' responses in switch and non-switch tasks.

Language task		Responses		
	Sum	Appropriate	% inappropriate	Omission
Non-switch tasks	464	441	18 (3.87)	5
Persian	232	215	14 (77.77)	3
English	232	226	4 (22.22)	2
Switch tasks	464	202	239 (51.50)	23
Persian-English	232	142	83 (36.82)	7
English-Persian	232	60	156 (65.27)	16

LT, Language task; Omission: responses scored as other, % inappropriate: the percentage of inappropriate responses (responses scored as "other" were not included).

TABLE 9 | Models of responses in Experiment 3.

Predictor	Estimate	SE	z-value	p
Language task as main predictor: $\chi^2(1) = 48.51$, $p < 0.001$, $N = 900$				
(Intercept)	−7.384	0.734	−10.060	< 0.001
Language task	3.811	0.420	9.078	< 0.001
Language task and target language as predictors: $\chi^2(1) = 83.747$, $p < 0.001$, $N = 900$				
(Intercept)	−4.820	0.538	−8.963	< 0.001
Language task	3.830	0.284	13.507	< 0.001
Target language	−1.737	0.201	−8.630	< 0.001

differed with respect to combinatorial processing in the three experiments. One of the main aims of the study was to discuss what might cause such differences, and what implications do the results of the present study have for language processing in bilingual speakers. I suggest that in the present study, different experimental contexts led to different patterns of control mechanism in bilingual language processing, because as Green (2011) states, differences in experimental contexts lead to differences in neural loci at which lexical items from the target language can be selected. Accordingly, all "speakers adjust their behavior during an experiment to the specific control demands it imposes" (Green and Abutalebi, 2013, p. 522). Different experimental contexts and the external instructions given to participants may lead to changes in the strength between the nodes within the network, suggesting that exogenous factors may affect the control mechanism (Green, 2011). To put it differently, the pattern of strength between the nodes (e.g., a lemma node and its corresponding combinatorial node) may vary depending on the context in which languages are used. Consequently, the changes in the strength between the nodes may yield in different linguistic behavior.
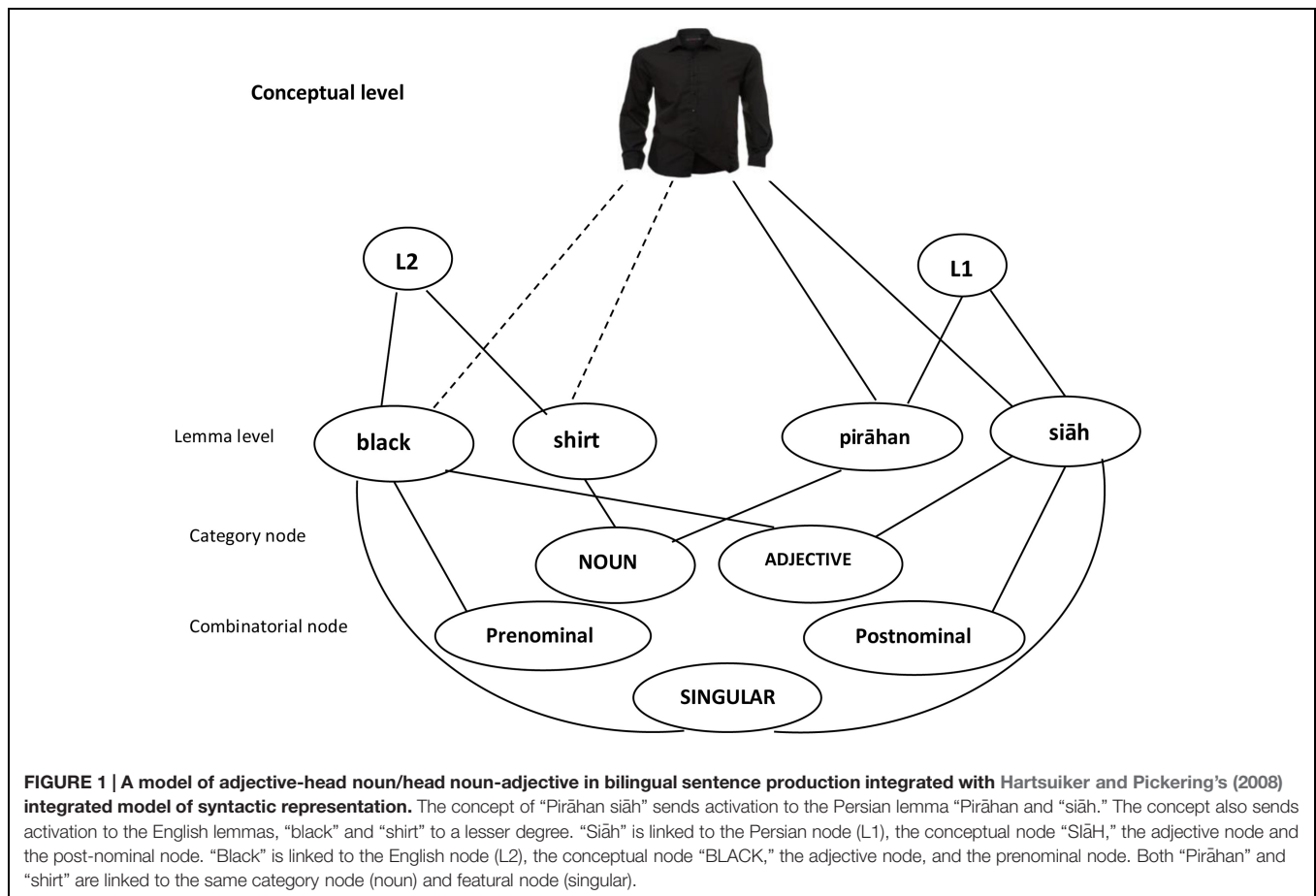
Now I consider how the results of the present study may be integrated with Hartsuiker and Pickering's (2008) integrated model of syntactic representation. Below an outline of the model is given first, followed by a description of the results using a model of adjective-head noun/head noun-adjective in bilingual sentence production. In the switch trials in Experiments 1 and 2, participants used both a noun and an adjective from the other language. In the non-switch trials, however, a noun and an adjective had to be selected from the base language. In the switch trials in Experiment 3, participants used only the adjective of the noun-adjective string from the other language. Thus, what is common in all experiments is that producing responses involves activating the appropriate noun lemma together with (a) its category information (noun), (b) its featural information (e.g., singular/plural), (c) the language node (e.g., Persian) and activating the appropriate adjective lemma together with (a) its category information (adjective), (b) its combinatorial information (prenominal/post-nominal), and (c) the language node (e.g., Persian). According to the model, when a Persian-English bilingual speaker intends to produce "pirāhan siāh" (lit. "shirt black"), the concept of "PIRAHAN SIAH" sends activation to the Persian lemma "pirāhan"(shirt) and "siāh"(black). Since the concept is shared between the two languages (Hartsuiker and Pickering, 2008),

it also sends activation to the English lemmas (i.e., "black" and "shirt") to a lesser degree (see Schoonbaert et al., 2007).

According to the model, "siāh" is linked to the Persian node (L1), the conceptual node "SIĀH," the adjective node, and the post-nominal node. "Black" is linked to the English node (L2), the conceptual node "BLACK," the adjective node, and the prenominal node (see **Figure 1**). Both "Pirāhan" and "shirt" are linked to the same category node (Noun). As stated above, when a Persian-English bilingual speaker intends to produce "siāh," first the conceptual node "SIĀH" is activated. Then activation spreads to the "siāh" lemma, the Persian language node, and the post-nominal node (combinatorial node). According to the model, the "SIĀH" conceptual node activates the "black" lemma as well, but since the "black" lemma receives little support from the language node (Persian), activation of the lemma "black"-belonging to the other language- is weaker (see Nicoladis, 2006). But even the little activation of the "black" lemma leads to the activation of the prenominal node to a lesser degree (Hatzidaki et al., 2011). In other words, while a Persian-English bilingual speaker normally uses the "siāh" adjective following a noun (i.e., he or she uses the post-nominal combinatorial node), sometimes he or she uses "siāh" before a noun (i.e., he or she uses the prenominal combinatorial node).

The results suggest that producing a Persian adjective, for instance, "derāz (long)" in a construction such as "xatkeš-e derāz" (lit. "ruler long"), causes the activation of the lemma node "derāz," the NA combinatorial node, the link between the lemma node (derāz) and the combinatorial node, and the category node (adjective). The combinatorial node retains activation at least temporarily (cf. Branigan et al., 1999), and a bilingual speaker is more likely to use the same combinatorial node (post-nominal) again even when using an English adjective (Cleland and Pickering, 2003). In other words, the concurrent activation of an NA combinatorial node might "lead to the strengthening of the link between the lemma nodes" (p. 217) in the other language of a bilingual (here English) and the NA combinatorial node. Cleland and Pickering (2003) suggested that "more generally, the activation of combinatorial nodes is related to the construction of constituent-structure representations" (p. 216). Accordingly, an NA combinatorial node is activated when a Persian adjective (e.g., qermez, meaning red) is used in the noun-adjective construction. In the same vein, an AN combinatorial node is activated when an English
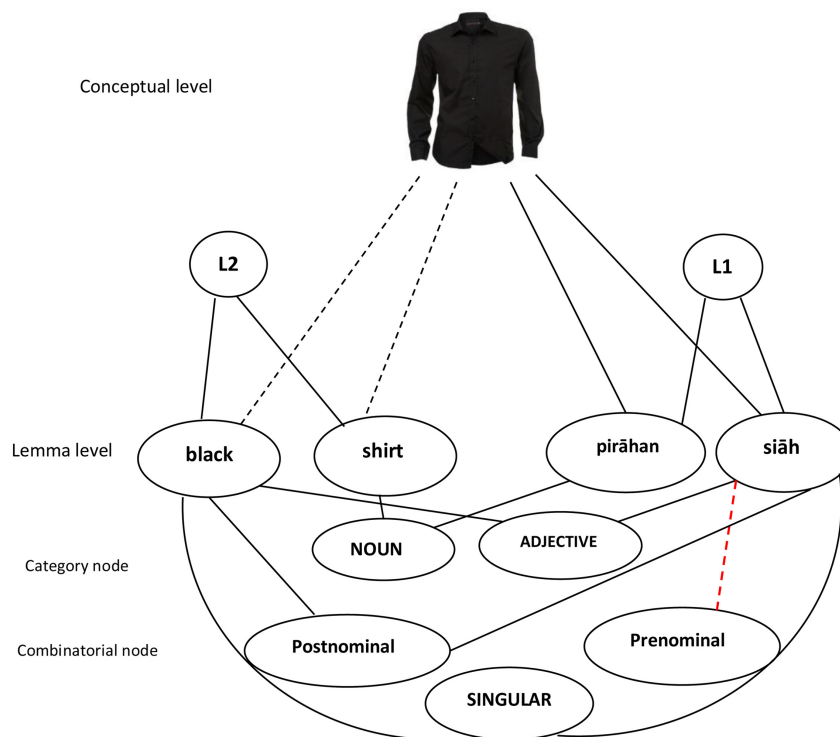
FIGURE 1 | A model of adjective-head noun/head noun-adjective in bilingual sentence production integrated with Hartsuiker and Pickering's (2008) **integrated model of syntactic representation.** The concept of "Pirāhan siāh" sends activation to the Persian lemma "Pirāhan and siāh." The concept also sends activation to the English lemmas, "black" and "shirt" to a lesser degree. "Siāh" is linked to the Persian node (L1), the conceptual node "SIāH," the adjective node and the post-nominal node. "Black" is linked to the English node (L2), the conceptual node "BLACK," the adjective node, and the prenominal node. Both "Pirāhan" and "shirt" are linked to the same category node (noun) and featural node (singular).

adjective (e.g., green) is used in the adjective–noun string (see Pickering and Branigan, 1998). Accordingly, producing an English construction involving an adjectival construction such as "long road" involves the prenominal adjectival modification, whereas producing a Persian NP such as "jāddeh-ye tulāni" (lit. "road long") involves the post-nominal adjectival modification. Thus, the constructions are associated with a combinatorial node, A,N and N,A nodes, respectively (Cleland and Pickering, 2003).

As stated above, since the link between a lemma node and a certain combinatorial node retains activated, it is more likely that the same link is used between a lemma node from the other language and the activated combinatorial node in the subsequent production of an adjective–noun string. This may explain why a Persian-English bilingual produces "mard-e old" (lit. "man old") after he/she produces a NA construction such as "māhi-e bozorg" lit. "fish big." Bilinguals' switching back and forth between the two languages has a critical role in increasing the activation of the non-target language lemmas and the syntactic information (i.e., featural and combinatorial information) associating with them. In the adjective case, this leads to using the combinatorial node from the other language (see **Figure 2**).

The results showed that 65, 80, and 78% of the responses in Experiments 1–3, respectively, in which participants used

the combinatorial node from the other language occurred after they used the same combinatorial node in the previous trail. The results demonstrated that participants had the tendency to produce sentences with the syntactic structure of a self-produced sentence during language production. Given this situation, producing a construction that employs a NA construction (e.g., "pesar-e mariz" lit. "boy sick") enhances the likelihood of producing the subsequent construction using the NA structure. This occurs because "combinatorial nodes retain activation after use" (Cleland and Pickering, 2003, p. 217).

There is a debate about whether grammatical feature selection is an automatic consequence of lexical node selection (see Schiller and Caramazza, 2003). Caramazza (1997) distinguishes between "intrinsic" grammatical features and "extrinsic" grammatical features. Intrinsic grammatical features are considered as inherent features of lexical items, however, extrinsic grammatical features are those features that "are not inherently associated with a word and are determined contextually (e.g., number, tense)" (Purmohammad, 2015, p. 88). Whereas 'gender' is considered as an arbitrary property, 'verb' is not an arbitrary feature of a lexical item. He suggests that the accessibility of different grammatical features is not uniform (Caramazza, 1997). Accordingly, while gender features are not automatically activated by the semantic network, tense and grammatical class (e.g., noun) features "do receive activation from the semantic network" (p. 195). I suggest

FIGURE 2 | A model of syntactic interference inside NP structure integrated with Hartsuiker and Pickering's (2008) **integrated model of syntactic representation.** The dotted red line shows the temporary link between a Persian adjective with the combinatorial node of the other language (English).

that as the combinatorial feature is an inherent feature of a lexical item, it is automatically activated. If this were the case, the question arises how we can account for cases where a word uses the combinatorial feature from the other language? According to the models that posit that inherent grammatical features are automatically activated (see Caramazza, 1997; Levelt et al., 1999) when an adjective (a lexical node) is activated, its combinatorial node (prenominal or post-nominal position) is automatically activated. If this were the case, I suggest that when an adjective is linked to the combinatorial node that belongs to the other language, its activated combinatorial node is suppressed (deactivated) and the syntactic feature from the other language is retrieved, instead. Thus, an additional locally control (i.e., a local reactive inhibition) is exerted in order to inhibit the activated syntactic feature (see Colzato et al., 2008 for the term "reactive"). If the Caramazza's (1997) account that when a lexical node it activated, its inherent grammatical features (e.g., the combinatorial node) are automatically activated were not the case, an alternative interpretation would be that a lexical node is directly linked to the combinatorial node that belongs to the other language without the need to suppress the word's intrinsic syntactic feature.

The results may also be interpreted in terms of the asymmetric switching cost account (see Meuter and Allport, 1999). In the present study (78, 89, and 65%) of the inappropriate responses in Experiments 1–3, respectively, occurred in switches from L2 to L1. The results are consistent with Meuter's

(1994) and Meuter and Allport (1999) study. Meuter and Allport (1999) reported that when a bilingual speaker switches, the cost of switching (reaction time) is greater when he switches from his L2 to his L1 than vice versa. In other words, switching in bilingual language production follows from asymmetric switching costs. The asymmetric switching cost account postulates that in code-switched utterances when the intended response language is participants' L1, we expect stronger recording of the distractor (see Meuter, 2005). Moreover, we expect more inappropriate responses when the intended response language is participants' L1. The results of the study are in line with Meuter (1994) and Meuter and Allport (1999) in that more responses (59%) scored as "other" occurred in switches from L2 to L1 suggesting that switches from L2 to L1 are more costly than vice versa. Participants had more difficulty making appropriate responses in switches from L2 to L1 than from L1 to L2, because bilingual speakers experience much more difficulty when they have to "suppress a resulting inappropriate response" (Meuter, 2005, p. 355) in their L1. According to Meuter and Allport (1999) the reason for the paradoxical pattern in the switch conditions is that the inhibition of L1 is considerably powerful in non-balanced bilingual speakers. Thus, the cost that arises from its removal is considerably large (see Green, 1993, 1998). To connect the Hartsuiker et al. (2004) model of syntactic representation model with Meuter and Allport's (1999) findings, the results of the present study reveal that participants had

more difficulty reactivating the combinatorial node (prenominal) of Persian when switching from L2 to L1. This yielded in more inappropriate responses in switches in this direction. Accordingly, the reason why less inappropriate responses were observed in switches from L1 to L2 may be that speaking in L1 requires little active inhibition of L2 (Meuter and Allport, 1999), therefore, in L1 to L2 switches participants needed less effort to reactivate their L2. Moreover, I interpreted the results of the present study in terms of the inhibitory processes (see Green, 1986, 1998). The presence of asymmetric language switch pattern is viewed as the main evidence supporting the use of inhibitory process (see Meuter and Allport, 1999; Costa and Santesteban, 2004). Thus, the results are in favor of the inhibition process in bilingual language production.

CONCLUSION

The results indicated that bilingual speakers may use a word from one language and the grammar from the other language. During bilingual language processing, the syntactic feature of a lexical item may undergo a local reactive inhibition and lexical items may use the syntactic feature from the other language, instead. As a combinatorial node of an adjective "retains activation at least temporarily" (Cleland and Pickering, 2003, p. 217), bilinguals are more likely to use the same combinatorial node again even when producing an adjective from the other language. The findings of the present study keep in line with the interference accounts of syntactic processing in bilinguals' language production, and the parallel activation of the two languages during speech production. More syntactic interference occurred in the switch tasks in which the two languages of a bilingual speaker were involved to a greater degree. Most of the inappropriate responses were produced in switches from L2 to L1 than from L1 to L2. While language proficiency did not put effects on responses, language task and target language significantly affected participants' responses.

REFERENCES

Aguirre, A. (1976). *Acceptability Judgments of Code-Switching Phrases by Chicanos: Some Preliminary Findings*, ed. ERIC. New York, NY: ERIC, 122, 129.

Arias, R., and Lakshmanan, U. (2005). "Code-switching in a Spanish-English bilingual child: a communication resource?," in *ISB4: Proceedings of the 4th International Symposium on Bilingualism*, eds J. Cohen, K. T. McAlister, K. Rolstad, and J. MacSwan (Somerville, MA: Cascadilla Press), 94–109.

Belazi, H., Rubin, E., and Toribio, A. J. (1994). Code switching and X-bar theory: the functional head constraint. *Linguist. Inq.* 25, 221–238.

Bernolet, S., Hartsuiker, R. J., and Pickering, M. J. (2007). Shared syntactic representations in bilinguals: evidence for the role of word-order repetition. *J. Exp. Psychol. Learn. Mem. Cogn.* 33, 931–949.

Branigan, H. P., Pickering, M. J., and Cleland, A. A. (1999). Syntactic priming in written production: evidence for rapid decay. *Psychon. Bull. Rev.* 6, 635–640. doi: 10.3758/BF03212972

Broersma, M. (2011). "Triggered code-switching: evidence from picture naming experiments," in *Modeling Bilingualism from Structure to Chaos: In Honor of Kees de Bot*, eds M. S. Schmid and W. Lowie (Amsterdam: John Benjamins), 37–57.

Bullock, B. E., and Toribio, A. J. (2009). "Themes in the study of code-switching," in *The Cambridge Handbook of Linguistic Code-Switching*, eds B. E. Bullock and A. J. Toribio (Cambridge: Cambridge University Press), 1–17.

Cantone, K. F., and MacSwan, J. (2009). "Adjectives and word order: a focus on Italian-German codeswitching," in *Multidisciplinary Approaches to Code Switching*, eds L. Isurin, D. Winford, and K. de Bot (Amsterdam: John Benjamins Publishing Company), 243–278.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cogn. Neuropsychol.* 14, 177–208. doi: 10.1080/026432997381664

Cleland, A. A., and Pickering, M. J. (2003). The use of lexical and syntactic information in language production: Evidence from the priming of noun-phrase structure. *J. Mem. Lang.* 49, 214–230. doi: 10.1016/S0749-596X(03)00060-3

Colzato, L. S., Bajo, M. T., van den Wildenberg, W., Paolieri, D., Nieuwenhuis, S., La Heij, W., et al. (2008). How does bilingualism improve executive control? A comparison of active and reactive inhibition mechanisms. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 302–312. doi: 10.1037/0278-7393.34.2.302

Costa, A. (2005). "Lexical access in bilingual production," in *The Handbook of Bilingualism. Psycholinguistics Approaches*, eds J. F. Kroll and A. M. B. De Groot (New York, NY: Oxford University Press), 309–325.

Costa, A., Roelstraete, B., and Hartsuiker, R. J. (2006). The lexical bias effect in bilingual speech production: evidence for feedback between lexical and sublexical levels across languages. *Psychon. Bull. Rev.* 13, 972–977. doi: 10.3758/BF03213911

Costa, A., and Santesteban, M. (2004). Lexical access in bilingual speech production: evidence from language switching in highly proficient bilinguals and L2 learners. *J. Mem. Lang.* 49, 491–511. doi: 10.1016/j.jml.2004.02.002

de Bot, K. (1992). A bilingual production model: levelt's 'Speaking' model adapted. *Appl. Linguist.* 13, 1–24.

de Bot, K., and Schreuder, R. (1993). "Word production and the bilingual lexicon," in *The Bilingual Lexicon*, eds R. Schreuder and B. Weltens (Amsterdam: Benjamins), 191–214.

Francis, W. S. (2005). "Bilingual semantic and conceptual representation," in *Handbook of Bilingualism: Psycholinguistic Approaches*, eds J. F. Kroll and A. M. de Groot (New York, NY: Oxford University Press), 251–268.

Gil, L. A., Eichler, N., Jansen, V., Patuto, M., and Müller, N. (2012). "The syntax of mixed DPs containing an adjective: evidence from bilingual german-romance (French, Italian, Spanish) children," in *Proceedings of the 14th Hispanic Linguistics Symposium*, eds K. Geeslin and M. Díaz-Campos (Somerville, MA: Cascadilla Proceedings Project), 242–257.

Green, D. W. (1986). Control, activation and resource. *Brain Lang.* 27, 210–223. doi: 10.1016/0093-934X(86)90016-7

Green, D. W. (1993). "Towards a model of L2 comprehension and production," in *The Bilingual Lexicon*, eds R. Schreuder and B. Weltens (Amsterdam: John Benjamins Publishing Company), 249–277.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Biling. Lang. Cogn.* 1, 67–81. doi: 10.1017/S1366728998000133

Green, D. W. (2011). Language control in different contexts: the behavioural ecology of bilingual speakers. *Front. Psychol.* 2:103. doi: 10.3389/fpsyg.2011.00103

Green, D. W., and Abutalebi, J. (2013). Language control in bilinguals: the adaptive control hypothesis. *J. Cogn. Psychol.* 25, 515–530. doi: 10.1080/20445911.2013.796377

Grosjean, F. (2008). *Studying Bilinguals*. Oxford: Oxford University Press.

Hamers, J. F., and Blanc, M. H. A. (1989). *Bilinguality and Bilingualism*, Revised edn, Cambridge: Cambridge University Press.

Hartsuiker, R. J., and Pickering, M. J. (2008). Language integration in bilingual sentence production. *Acta Psychol.* 128, 479–489. doi: 10.1016/j.actpsy.2007.08.005

Hartsuiker, R. J., Pickering, M. J., and Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychol. Sci.* 15, 409–414. doi: 10.1111/j.0956-7976.2004.00693.x

Hatzidaki, A., Branigan, H., and Pickering, M. J. (2011). Co-activation of syntax in bilingual language production. *Cogn. Psychol.* 62, 123–150. doi: 10.1016/j.cogpsych.2010.10.002

Karousou-Fokas, R., and Garman, M. (2001). Psycholinguistic interpretation of codeswitching: evidence from fluent Greek–English bilingual adults. *Int. J. Biling.* 1, 39–69. doi: 10.1177/13670069010050010301

Kootstra, G. J., van Hell, J. G., and Dijkstra, T. (2010). Syntactic alignment and shared word order in code-switched sentence production: evidence from bilingual monologue and dialogue. *J. Mem. Lang.* 63, 210–231. doi: 10.1016/j.jml.2010.03.006

Kroll, J. F., Bobb, S. C., Misra, M., and Gue, T. (2008). Language selection in bilingual speech: evidence for inhibitory process. *Acta Psychol.* 128, 416–430. doi: 10.1016/j.actpsy.2008.02.001

Kroll, J. F., Bobb, S. C., and Wodnecka, Z. (2006). Language selectivity is the exception, not the role: arguments against a fixed locus of language selection in bilingual speech. *Lang. Cogn.* 9, 119–135. doi: 10.1017/S1366728906002483

La Heij, W. (2005). "Selection processes in monolingual and bilingual lexical access," in *Handbook of Bilingualism: Psycholinguistic Approaches*, eds J. F. Kroll and A. M. de Groot (New York, NY: Oxford University Press), 289–307.

Levelt, W. J. M., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behav. Brain Sci.* 22, 1–75. doi: 10.1017/S0140525X99001776

MacSwan, J. (2009). "Generative approaches to code-switching," in *The Cambridge Handbook of Linguistic Code-Switching*, eds B. E. Bullock and A. J. Toribio (Cambridge: Cambridge University Press), 309–335.

Mahootian, S. (2006). "Code switching and mixing," in *Encyclopedia of Language and Linguistics*, 2nd Edn, ed. K. Brown (Oxford: Elsevier).

McClure, E. (1981). "Formal and functional aspects of the codeswitched discourse of bilingual children," in *Latino Language and Communication Behavior*, ed. R. P. Duran (Norwood, NJ: ABLEX Publishing Corporation), 69–94.

McClure, E. (1977). "Aspects of code-switching in the discourse of bilingual Mexican-American children," in *Georgetown University Round Table*, ed. M. Saville-Troike (Washington, DC: Georgetown University Press).

Meuter, R. F. I. (1994). *Language Switching in Naming Tasks*. Oxford: University of Oxford.

Meuter, R. F. I. (2005). "Language selection in bilinguals: mechanisms and processes," in *Handbook of Bilingualism: Psycholinguistic Approaches*, eds J. F. Kroll and A. M. de Groot (New York, NY: Oxford University Press), 349–370.

Meuter, R. F. I., and Allport, A. (1999). Bilingual language switching in naming: asymmetrical costs of language selection. *J. Mem. Lang.* 40, 25–40. doi: 10.1037/a0034060

Myers-Scotton, C. (1993). *Duelling Languages: Grammatical Structure in Codeswitching*. Oxford: Clarendon Press.

Myers-Scotton, C. (2002). *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*. New York, NY: Oxford University Press.

Nartey, J. (1982). Code-switching, interference or faddism? Language use among educated Ghanians. *Anthropol. Linguist.* 24, 183–192.

Nicoladis, E. (2006). Cross-linguistic transfer in adjective-noun strings by preschool bilingual children. *Biling. Lang. Cogn.* 9, 15–32. doi: 10.1017/S136672890500235X

Pandit, I. (1990). "Grammaticality in code switching," in *Codeswitching as a Worldwide Phenomenon*, ed. R. Jacobson (New York, NY: Peter Lang), 33–69.

Paradis, M. (1993). Linguistic, psycholinguistic, and neurolinguistic aspects of interference in bilingual speakers: the activation threshold hypotheses. *Int. J. Psychol.* 9, 133–145.

Pfaff, C. W. (1979). Constraints on language mixing: intrasentential codeswitching and borrowing in Spanish/English. *Language* 55, 291–318. doi: 10.2307/412586

Pickering, M. J., and Branigan, H. P. (1998). The representation of verbs: evidence from syntactic priming in language production. *J. Mem. Lang.* 39, 633–651. doi: 10.1006/jmla.1998.2592

Poplack, S. (1980). Sometimes I'll start a sentence in Spanish y termino en espanol: towards a typology of codeswitching. *Linguistics* 18, 581–618. doi: 10.1515/ling.1980.18.7-8.581

Purmohammad, M. (2015). *Codeswitching: A Touchstone of Models of Bilingual Language Production*. Ph.D. thesis, University of Bern. Switzerland.

Sankoff, D., and Poplack, S. (1981). A formal grammar for codeswitching. *Papers Linguist.* 14, 3–46. doi: 10.1080/08351818109370523

Schiller, N. O., and Caramazza, A. (2003). Grammatical feature selection in noun phrase production: evidence from German and Dutch. *J. Mem. Lang.* 48, 169–194. doi: 10.1016/S0749-596X(02)00508-9

Schoonbaert, S., Hartsuiker, R. J., and Pickering, M. J. (2007). The representation of lexical and syntactic information in bilinguals: evidence from syntactic priming. *J. Mem. Lang.* 56, 153–171. doi: 10.1016/j.jml.2006.10.002

Schwartz, A. I., and Kroll, J. F. (2006). "Language processing in bilingual speakers," in *Handbook of Psycholinguistics*, 2nd Edn, eds M. J. Traxler and M. A. Gernsbacher (Amsterdam: Elsevier), 967–999.

Selles, A. (2011). *Cross-Linguistic Interference in the Production of Adjective and Noun Word Order by Spanish-English Bilinguals*. Master's theses, University of Edinburgh, Edinburgh.

Sobin, N. J. (1984). On code-switching inside NP. *Appl. Psychol.* 5, 293–303. doi: 10.1017/S0142716400005191

Voga, M., and Grainger, J. (2007). Cognate status and cross-script translation priming. *Mem. Cogn.* 35, 938–952. doi: 10.3758/BF03193467

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

CrossMark

Multiple Grammars and the Logic of Learnability in Second Language Acquisition

*Tom W. Roeper** *

Department of Linguistics, University of Massachusetts, Amherst, Amherst, MA, USA

The core notion of modern Universal Grammar is that language ability requires abstract representation in terms of hierarchy, movement operations, abstract features on words, and fixed mapping to meaning. These mental structures are a step toward integrating representational knowledge of all kinds into a larger model of cognitive psychology. Examining first and second language at once provides clues as to how abstractly we should represent this knowledge. The abstract nature of grammar allows both the formulation of many grammars and the possibility that a rule of one grammar could apply to another grammar. We argue that every language contains Multiple Grammars which may reflect different language families. We develop numerous examples of how the same abstract rules can apply in various languages and develop a theory of how language modules (case-marking, topicalization, and quantification) interact to predict L2 acquisition paths. In particular we show in depth how Germanic Verb-second operations, based on Verb-final structure, can apply in English. The argument is built around how and where V2 from German can apply in English, seeking to explain the crucial contrast: "nothing" yelled out Bill/*"nothing" yelled Bill out in terms of the necessary abstractness of the V2 rule.

Keywords: multiple grammars, learnability, transfer, acceptability/grammaticality judgments, minimalism, verb-second, interfaces

INTRODUCTION[1]

Modern Minimalism (Chomsky, 1995, 2013) has made grammatical description dramatically more abstract. If it is on the right track, then it should reflect typical scientific progress: it should both transparently capture deeper rules and make more refined predictions about grammatical detail (Adger and Smith, 2005).

We argue that Minimalism naturally forces an approach to every speaker's grammatical knowledge in terms of Multiple Grammars (MG), which in turn enables sharp predictions about second language acquisition (L2) to arise. Thus, another virtue of modern minimalism is that it invites a new domain of data to be directly relevant to linguistic theory—and, we may reasonably expect, to pedagogical practice.

Unlike, for instance, Phrase Structure Rules, which are stated in language particular terms, we can state rules in terms of Merge, Labeling, Feature-satisfaction. When the rules are stated as abstractions in terms of abstract minimally labeled categories, they immediately apply to many languages. It follows that a speaker of L1 can—and I argue must—apply these rules to L2.

---

[1] Thanks to Luiz Amaral, Leah Bauke, David Erschler, Stefan Keine, Terje Lohndal, Tom Rankin, Marit Westergaard, Rebecca Woods, the class on Variation with Lisa Green, and Frontiers reviewers for helpful discussions.

This perspective leads to a little logic which is the backbone of this essay (Roeper, 2011, 2014):

(1) a. Grammars must be simple. They are not encumbered with notation for choice and optionality.

   b. Principles of grammar are very abstract.

   c. Such abstract grammars will inevitably apply to many languages, not just one.

   d. All speakers will possess multiple rules, in effect Multiple Grammars, which must automatically be applicable in a second language.

   e. If true, then the study of L2 acquisition (and variation in general) provides a unique domain to test whether these levels of abstraction are appropriate.

Language typology suggests that Universal Grammar (UG) captures a number of language families with simple alternations: VO or OV, Wh-movement, or non-wh-movement, Scope inversion or surface scope, Pragmatic Object deletion, or not (Hot/Cool distinction) It is natural to expect that every speaker will have all prototypes available—especially, but not exclusively, if they speak an L2 that carries them.

This approach, in turn, can be viewed as a version of Transfer (Full Access, Full Transfer), common in the literature and assumed by many (see White, 2003 for an overview). However, it does not entail that a rule of one grammar is actually transferred into another grammar. Why not? Because the simplicity of rules would not allow a rule to be stated with all the contingencies that are found in every particular grammar. For any particular language, many specific lexical contingencies and structural variations are involved, as we shall detail, and they exhibit the over-application and under-application of rules that is typical of the variation found in L2.

In sum, MG theory follows from both the spirit and substance of the Minimalist approach. Fortunately, the grammatical subtlety of work in L2 has matured to the point where we can begin to develop a Learnability Logic and predict cross-linguistic L2 acceptability judgments under the assumption that more than one grammar is simultaneously available[2]. The goal here will be to show why and how such logic should work, examine cases where it applies, and predict L2 acceptability/grammaticality judgments. In other words, to seek theoretically predictable "grammaticality" from an L2 perspective in the same manner we seek to predict ungrammaticality in typical theoretical work.

Many of our arguments are implicitly present in a number of recent papers that advocate Full Access/Full Transfer, Feature-Reassembly, and Variational Learning (for which MG theory is a prerequisite). Our goal is to take a few steps—just a few—toward the technical precision found in monolingual generative studies.

---

[2]This essay is written by someone who works primarily in L1 acquisition. It is important that discussions which cross boundaries between close disciplines are promoted. One consequence is that it is virtually impossible to be familiar with the whole voluminous L2 literature and therefore relevant arguments and evidence may exist of which I am unaware or unable to explore carefully. I would be happy to receive any information relevant to both the theoretical and empirical claims advanced here by those with more detailed knowledge. The broad focus on FA/FT, Feature-Reassembly, and Variational Learning I take to be a fair sample of currently explored approaches, but I welcome further detailed information pertinent to the claims in this paper, which may have easily escaped my attention.

Overview

The essay is focused on a detailed discussion of an abstract version of V2, formulated with open XP and YP environments:

$$V2\ Rule: XP\ YP\ V => XP\ V\ YP.^{3}$$
$$1\quad 2\ 3\qquad 1\quad 3\ 2$$

This leads to predictable overapplication of rules in Wh-movement, Quotation, Topicalization, particle behavior, and potential interactions with Information Structure, which can then influence L2. This in turn may yield greater insight into V2.

Section Avoiding the Concept of Transfer considers Dimensions of Transfer and Compatibility among Modules. Section Diverse V2 considers V2 in terms of different Force types (Wh-movement, Quotation, Topicalization) and an extensive discussion of particles. Section Moving Toward L2 Formalism contrasts V2 I L1 and L2 environments, arguing that the acquisition path of L1 reveals the right level of abstraction. Then in Section Moving toward L2 Formalism we consider Information Structure and Discourse, and the Principle of Minimal Modular Contrast, and the role of Exhaustivity in Topicalization. The next Section considers Missing Subjects and Objects, and then considers Dialects and Modular Incompatibility. Then we return to a more extensive discussion of and integration with MG theory with Variational Learning, Full Transfer and Full Access, and Feature Re-assembly. Section Conclusion, concludes the essay with a summary of why Minimalism makes particular predictions about L2.

AVOIDING THE CONCEPT OF TRANSFER

The term "Transfer" has been pivotal in much of the discussion but in my view it both overstates and understates how L1 influences L2. In one sense, the MG approach builds on the core intuition of Transfer, but in another sense our goal is to argue against Transfer as the core concept and dissolve the term into concepts that reach further into making predictions based on UG for L2 acquisition. If separate grammars are simultaneously present in speakers, the status of these two grammars will vary with how far advanced L2 knowledge has developed. And it may affect comprehension and production quite differently[4].

Transfer Dimensions

What is <u>Full Transfer</u>? Although the term is widely used, it is not always clear what the status is of something transferred in

---

[3]We use the term "rule" for convenience, without prejudice against the view that rules can be reduced to principles. And we use the term Grammar in Multiple Grammars to refer to something as narrow as a single rule if it reflects a core property of a language family. We use the term "module" in the traditional sense for Binding, wh-move, thematic roles.

[4]FT/FA hypothesizes that the initial state of L2 acquisition is the final state of L1 acquisition (Full Transfer) and that failure to assign a representation to input data will force subsequent restructurings, drawing from options of UG (Full Access). By contrast, our approach suggests that simple rules allow the grammars to remain present as distinct entities, producing unexpected ambiguities. The idea of Minimal trees suggested by Vainikka and Young-Scholten (1996) is very much in the spirit of seeking the minimalist substructure that can transfer, but we do not argue that more refined features cannot have an impact.

each respective language. One extremity seems to be: If one fully Transfers part of one grammar into another, then it is a compatible and unrecognizable part of that language. In effect, one language incorporates or copies part of another. That is true of the many words borrowed from one grammar into another. Speakers are often unaware of their origins. They are <u>not</u> present in a second grammar in the speaker's mind. <u>Charm</u> came from French, but is pronounced by English rules, and it does not entail the presence of a French grammar in an English speaker.

In fact Full Transfer would entail many special statements with respect to other modules: case, agreement, thematic roles, wh-movement, LF movement. The requirement on Simple Rules aims to avoid exactly this complexity. We will advance a general proposal that limits the interaction of modules below called: <u>Minimal</u> <u>Modular</u> <u>Contact</u>. It, obviously, remains a programmatic goal, but it can guide research nonetheless.

Transfer Induced Ambiguities and Modular Compatibility

What is a paradigm case of Transfer? Consider (2)

(2) the dog chased the cat

It is either SVO or OVS (via V2 on SOV) in several Germanic languages, but only SVO in English. Yet speakers of those Germanic languages demonstrably (as we discuss below) register the other meaning. Under a Multiple Grammars (MG) approach[5] the argument is not simply that L1 is used in an L2 environment. Such examples obey what we can call:

(3) Complete Modular Compatibility

That is, no other module of L2 is disobeyed as in (2)[6]. So (3) predicts that for instance, V2 can be applied in sentences to create OVS structure just in case, the Subject-verb AGREEMENT module is not violated, as it would be below (4):

(4) <u>the</u> <u>dogs</u> <u>are</u> chasing the cat => *dogs = object.

When this criterion is imperfectly observed, as it is sometimes, then we predict degrees of acceptability of MG in L2 (an issue raised by a reviewer) depending upon the status of the Agreement module in a grammar. This occurs when there is a variable state of L1 and L2 within individual speakers. How could there be a variable status of a module? German, for instance, has a very complex two-layered system of case-assignment and agreement. If one had only the Strong system or only some of the Agreement paradigm, then the appeal of a rule without agreement is stronger.

So we predict, if favored by pragmatics, a first stage German or Dutch learner of English (which happens, see below) might indeed allow: <u>the</u> <u>rats</u> <u>are</u> <u>chasing</u> <u>the</u> <u>dog</u> to mean <u>the</u> <u>dogs</u> <u>chase</u> <u>the</u> <u>rats</u>, while the advanced learner, as evidence below suggests, will inevitably continue to allow only (2) to be misanalyzed in comprehension when Subject-verb agreement is unmarked.

Thus, it captures what can be descriptively observed as a moment of L1/L2 Non-interference:

(5) L1/L2 Non-interference:
     Rule X from L1 can apply in L2 because

     a. no L2 module is violated.
     b. no obligatory L1 module is ignored.

This is a perfect case where we argue that nothing has been "Transferred," but rather <u>L1</u> <u>simply</u> <u>operates</u> <u>in</u> <u>L2</u>. The same holds for the application of Inverse Scope for quantification in an L2 with only Surface Scope where, for instance, the Case-assignment module would make no difference. Similarly, if the Interface with Logical Form is universal, then it does not have to be learned or Transferred, which we discuss. Nonetheless, we argue that within L2, L1/L2 non-interference can make surprisingly subtle predictions. We show where the grammar of separable particle constructions in German could be involved in the L2 analysis of English particle behavior in quotation.

In sum, MG theory is a significant form of grammatical economy: information is not written twice, but rather a single representation is accessible in more than one language. We expect that every grammar will contain simple pieces of UG that are not exclusive to it, but linked to another grammar type, hence Bilingualism is universal[7].

The critical reader might point out that this approach is akin to the claim that there are Linguistic Universals that can be innate and therefore do not have to be separately stated. Indeed MG theory is really an extension of traditional UG assumptions.

Is Genie Relevant?

One intriguing question can be asked (as a reviewer did): does one access UG through L1? The answer should be "no" because a speaker should still have access to UG as a set of inborn options, such that he could, for instance, set the pro-drop parameter the opposite way. However, it does seem that L1 is necessary to trigger the availability of UG under the Critical Period hypothesis.

If that were not the case, then Genie (Curtiss et al., 1974), the 12 year-old child discovered without a first language, would have learned English just the way that any L2 learner of English would in high school. But she was unable to. So we would argue that L1 is necessary to trigger UG, but not that one must "go through" L1 to locate UG or basic rules.

DIVERSE V2

We will now examine a series of less frequent constructions in English where V2 might apply.

In "Universal Bilingualism" (1999), I argued that English Quotation should not be collapsed with other forms of inversion into a complex rule (such as the version in Collins and Branigan, 1997) but remain a UG option selected by both English and German in a form that remains simple, following a basic principle of Avoid Complex Rules (see Amaral and Roeper, 2014 for more specifics). We will examine among other constructions:

(6) Wh-movement
     Quotation

---

[5]In Roeper (1999) and Amaral and Roeper (2014) for MG and L2.

[6]We take the term module to include the traditional ones: Case, Agreement, Move-alpha, LF-formation, although some recent theories suggest that modules can be replaced by interface requirements, this seems premature.

[7]See Lohndal (2013) for background discussion.

Topicalization
Empty Objects
Auxiliary inversion

V2 and Wh-Questions

Could V2 be present in Wh-questions? Rankin (2015) suggests that L1 may be *covertly* applied and rejected unconsciously quite frequently in the process of comprehension. This leads to unseen labor from L2 speakers in many grammatical situations. These are often assumed to be "processing" difficulties, but I consider them instances of representational conflict between two grammars. The comprehension dimension may exist even when a Speaker knows to avoid V2 in production.

Rankin (pc) suggests, based on judgment data, that there appears to be a difference in how L1 German learners of English comprehend these two sentences (in ongoing work):

(7) a. who woke up John
     b. who woke John up.

German speakers allow who in (7b) to have an object reading (John woke who up) while in (7a) who has a subject reading (someone woke up John). Both sentences involve movement of the Main Verb, while one moves the particle as well (7a). Only (7a) receives a purely English analysis (John = object), while (b) allows post-verbal John to be subject.

The implicit reasoning seems to be this: German disallows a particle to move with its verb (8d), while it is optional in English (8a,b):

(8) a. John woke up Bill
     b. John woke Bill up
     c. Hanns weckt ihn auf.
          [John woke him up]
     d. *Hanns weckt-auf/aufweckt ihn.
          [John woke up/upwoke him]

Therefore, the fact that verb+particle moves as a whole (8a), signals that English grammar is used, because that is impossible in German. Consequently movement is to the TP, leaving the subject in SPEC, which guarantees that when who is moved to CP, it remains a subject in (7a). Had the object moved in English, then Tense would move to C where it would have forced do-insertion:

(9) who did John wake up?

But if only the Head verb (wake) and the wh-word both move to CP, as in German (8c), then John could be the subject left in IP in who woke John up leading to the evident miscomprehension.

The unseparated movement of verb+particle blocks access to V2 from the German-L1 speaker's perspective. There is no verb in English *to upwake, therefore it must be movement to the English permitted position. Once again, the other case (10) meets the Non-interference Identity criterion:

(10) who woke John up?

(10) could be either (a) movement of wake to T or (b) movement to C, with no other modules disturbed in *either* case. Hence the German speaker cannot restrain

the V2 reading in comprehension. These results hold for advanced speakers, Rankin argues, and therefore are not typical production errors, but they show the continued presence of MG.

Quotation

From another angle, exactly the same contrast is at work in quotation as Amaral and Roeper (2014) show based on Bruening (2013), Collins and Branigan (1997), and Alexiadou and Anagnostopoulou (2007). If V2 applies in English, as (11) suggests, then it will allow the verb+particle to move to C:

(11) a. "nothing" yelled out John

and no do-insertion is allowed:

     b. *"nothing" did John yell out.

Of greater interest here, is the fact that in English the particle cannot be left behind (12a) while in German it must be left behind (12c):

(12) a. *"nothing" yelled John out.
     b. * "nichts" ruft aus/ausruft Hanns.
          ["nothing" yelled-out/out-yelled John]
     c. "nichts" ruft Hanns aus.
          ["nothing" yelled John out]

The German facts are not surprising and lead to a prediction. If questioned, German speakers should find (11b) acceptable (which my informal exploration supports) because only the Main Verb, that is the Head, not the particle can move under V2 in German, which is in line with other constraints on Head-movement.

It is important to note that only verb+particle moves over the subject. It is not the case that any larger VP (said to Bill, screams angrily) can move over the subject in English as these data show:

(13) a. *"no" said to Bill John
     b. *? "yes indeed" screamed angrily Fred

Therefore, the constraint to Head-movement which characterizes V-to-I movement in many grammars is upheld. The fact that verb+particle must move to C in quotation, using V2 and disallowing a stranded particle, calls for a more refined interpretation of the interaction of V2 and separable particle verbs. If the use of V2 in English is only stateable at an abstract level, then it will not be able to access the separable property of particles in the English lexicon, and therefore the whole verb+particle moves. If true, we might expect to find the restriction elsewhere. We turn now to other restrictions on particle movement which enlarge the question.

Topicalization

While German allows Topicalization of almost any element, precisely particles are excluded in both English and German?

(14) a. English: *out spread John, *over came Bill
          *out cried he, *up threw Mary, *on carried Fred

b. German: *Aus ruft er.
   [Out yelled he.]
   "He yelled out"[8]

Why? It is noteworthy that there are subtle differences here which point to the question of what the exact Label on the particle is:

(15)   a. *über läuft er [*over ran John]
       b. 'rüber läuft er [here over ran John]

(15b) is a reduced form of an adverb phrase <u>herüber</u> <u>läuft</u> <u>er</u> (here-over ran he), via a morphological rules that adds <u>her-</u>.
   The same holds for English with stylistic inversion:

(16)   over here came Bill.

and we would predict that these judgments would hold across L2-English and L1-German, but they deserve examination.

   These judgments obey the underlying abstraction that, while verb movement involves just Heads, movement to a Topic position must involve a phrase, a Maximal Projection. The Label on a particle is not absolutely clear, but it is evidently not projectable to a Maximal Projection.

Particle Mystery

A deeper question is why we <u>can</u> leave the particle behind with movement to TP:

(17)   John yelled the answer out.
       [likewise: shout out/scream out/holler out/bellow out]

but not with movement to C in English?[9] Interestingly Quotation does not allow particle-stranding either, even when movement is clearly to TP (18a):

(18)   a. * John hollered "I can't come" out.
       b. John hollered out "I can't come"[10] .

And this holds for passive as well:

(19)   *"I can't come" was hollered by John out

although one can have the by-phrase move over indirect objects (20a), just not the particle (20b):

(20)   a. Presents were given by Bill to John
       b. *Presents were given by Bill out.
       c. Presents were given out by Bill

And this same constraint applies to expletive cases like:

(21)   a. there walked a man in/ *there came a man over.
       b. there <u>walked</u> <u>in</u> a man/there came over a man.

---

[8]Leah Bauke (pc) points out that inherent reflexives block Topicalization as well:

i.   * Sich hat Hanns mit Maria gestritten.

This falls into line as well as an example which fails to be a Maximal Projection or have an acceptable Label.
[9]Müller (2003) is an edited volume where a common theme is that CP is a "vulnerable domain" in L2 and in disorders. This also seems to suggest that that it is linked to an Interface and therefore, perhaps, more easily subject to alteration or error. In the case at hand, it would have to explain the opposite—a tighter requirement on verb-raising which involve necessarily moving the particle as well.
[10]It is not linked to length. A monosyllable is also excluded: * I hollered "no" out.

And, finally, an old puzzle about particles remains:

(22)   a. John threw Bill <u>down</u> an apple
       b. *John threw Bill an apple <u>down</u>.
       c. he read me <u>out</u> the riot act
       d. *he read me the riot act <u>out</u>.

(23)   a. He tossed up these apples to the kids, but he could not toss them up those.
       b. . . . *toss them those up.

In (23b) there is no hint of Heavy-NP shift at work. We conclude that something much deeper, still unknown, is at work here.

MOVING TOWARD L2 FORMALISM

An important question buried in the discussion is this: if we have access to L1 in L2, then when is it operative? We have argued, as others have (Westergaard, 2003) that when full Non-interference Compatibility is present, then the application of L1 is unstoppable at the comprehension level. Where refined features of a structural description must be accessed—then the application is less automatic and less under speaker control, as in the examples just mentioned.

   This can be discussed in somewhat more formal terms. Whatever allows the intermediate appearance of a particle must be stated over a full Structural description (to use older terminology, Chomsky, 1957):

(24)   NP2 NP1 verb+particle => verb NP2 NP1 <u>particle</u> ⇒
                      verb NP2 <u>particle</u> NP1 *trace*.

Because we have a rule of Heavy-NP-shift one might imagine that it is a subpart of that rule which is expressed specifically when particles are present.

   One might also advance the view that a local phonological operation is involved, so that it is very narrow inversion. But exactly this suggestion would not explain why one cannot move over a simple subject:

(25)   * "no" screamed Bill out.

Altogether these facts point to the conclusion that it is only over a single <u>object</u> that one can move a particle verb[11].

   One reason that I explore this mystery here is that the behavior of L2-speakers might easily supply clues about the right level of abstraction that is relevant for the L1 description. In fact, numerous controversies over the ideal rule in a given language might be resolved if we treat the L2 data as pertinent to the original grammatical description, rather than assuming that L2 exploration should only proceed where the L1 is

---

[11]All the corners should be explored. Consider incorporation which also blocks incorporation from a stranded particle (ii) stranding:

i.   ??paper-up-picking
ii. ** up-paper-picking

See Bauke and Roeper (2011) for a morphological analysis in terms of Labeling theory. One might predict that a German L1 speaker would find (i) more acceptable because particle verbs in German allow incorporation.

well-understood.[12] . Will German L2 speakers of English accept or reject this entire array of facts about stranded particles? If they apply German V2 to English, then the first hypothesis is that all the stranded particle sentences should be acceptable. It is difficult to reason further until we have the evidence.

Since this domain is ripe for seeking L2 judgments, let us make some predictions about judgments of Germans learning English:

(26) 1. German speakers would find it fine to say:

a. (20b) "nothing" yelled John out.

because they invoke the requirement of separability for prefixes.
They will easily comprehend and produce:

b. "Nothing" said Bill

which has no particle and is pure V2. They will comprehend, as extended V2, but perhaps resist producing:

c. "nothing" yelled out Bill.

because the particle should be and can be left behind in German.

2. Both English and German speakers will reject:

a. *over came Bill

because neither language allows Topicalization of a particle, but both will accept:

b. over here came Bill

because it has been Relabeled in the lexicon as an Adverb. In this case, we predict that the idiomatic form in English (which violates the constraints):

c. in walked John

will be rejected by German speakers because it is treated as a separable prefix (eingehen).

How does (c) become acceptable in English? Suppose we argue that it can undergo Relabeling as an Adjunct Adverb in English, which is impossible without a further morpheme in German (ein => herein). Is there a theoretical implication here? Is idiomaticity the full explanation here? It could be that Relabeling in the lexicon from particle to adverb requires additional morphology in German but not English. Therefore, the German speaker cannot relabel the bare particle in, because it would require Relabeling via a morpheme, which is available

in the lexicon, but cannot be accessed in the syntax. A deeper principle could be involved: L1 cannot apply to L2 if it would entail a rule that crosses an Interface[13] . Another prediction is that all who reject (b) will also reject (c).

Further predictions:

3. German speakers will allow double objects in English

a. *he read me the book out

since analogous forms are acceptable in German

b. "Er hält mir nie die Tür auf" (Google)
                       [he holds me never the door open]

Now we can speculate about why it should be acceptable. In English, following Keyser and Roeper (1992), the dative and particle occupy the same position, which blocks even isolated dative objects:

b) he yelled the answer out to me.
       *?he yelled me out the answer.

while in German a productive benefactive dative pronoun exists which could be used in English (also available in English dialects). This makes the further prediction that a full nounphrase might be blocked for the German speaker:

(27) ?He yelled John the answer out.

One might regard this as a highly local point of "transfer." Instead we can sketch the following argument: suppose that using the German benefactive dative in English is a late adjunction in the derivation of the sentence and therefore it satisfies the Compatibility criterion.

If German speakers "accepted" this sentence, but never would use it, then it becomes a clear example of allowing an alternate German grammar to operate in English, not a "transfer."

These predictions call for careful study, but the cross-linguistic reasoning should be clear.

The upshot here is again that verb+particle behavior remains an unsolved mystery[14] . The larger array of data suggests that the simple V2 rule is involved with a further constraint consistent with the formulation of abstract rules:

(28) Apply rule to the highest possible Projection

This would favor not separating verb and particle and therefore taking a V node that would not see a division between verb and particle.

L2 Path and the L1 Path

One angle from which to examine the formalism is to ask:
    Does L2 acquisition mirror the L1 path?
    Amaral and Roeper (2014) argued that V2 is acquired piecemeal in L1, resolving a disagreement between Wexler (1998) and Yang (2002). Wexler (2011) argued that children acquired V2 very early because forms like:

---

[12]As I wrote this article I realized that I was carrying out an implicit experiment on myself. My L2 knowledge of German is regarded as very good and I trust myself to create examples. However, when seeking to establish that this constraint did not apply to German, I started to feel uncertain about (ii), although I suspect it would earn a strong ** from a native speaker and not a "?," but I cannot tell:

i. Er liest mir das Buch aus
ii. ?Er liest mir aus das Buch

One might presume that it is contemplation as a linguist which is the source of my uncertainty, but I suspect that it would be widespread and would deserve a careful study together with all the facts mentioned here (see Rankin, 2013 and references therein).

[13]Thanks to Leah Bauke for discussion.

[14]Bruening (2013) suggests that phonological parallelism constraint is present to capture facts about quotation, but it does not deal with facts like these which clearly belong to the puzzle.

(29) da geht er [there goes he]

occurred very early in the 2 year range, while Yang argued that V2 was not acquired until very late because object movement did not arise until 4 years[15] in children, nor was it frequent in the input:

(30) Fleisch isst er [meat eats he]

In the adult grammar, V2 is expressed with respect to an abstract maximal projection, XP. Evidence of late OVS from Yang but early LOC-V-S suggests that children proceed through a series of specified XP's before they generate the full abstract rule, which no one to my knowledge has thus far carefully traced:

(31) Typical forms of V2:

    a. Subject NP: **Er** frisst Fleisch.
He eats meat.
    b. LOC: **Da** sing er.
There sings he.
    c. ADV: **Schnell** fährt er.
Quickly moves he.
    d. DO: **Fleisch** frisst er.
Meat eats he.

(32) Less discussed forms of V2:

    a. Quotation: **"Willkommen"** sagt Herr Anders.
              "Welcome" said Mr Anders.
    b. VP fronting: **Zu mir allein kommen** will er nicht.
              To me alone come wants he not.
              "He does not want to come alone to me."
    c. Empty Topic in Discourse:
              Wo ist das Fleisch? __ frisst Hanns schon.
              Where is the meat? __ ate John already.
    d. Conjunction:
        Hanns spielt oft, **so** kann er ohne Mühe uns helfen.
        Hanns plays often, so can he without difficulty us help.
        "Hanns plays often, so he can help us without difficulty."

Left-Dislocation Option

One piece of evidence from Roeper (1973) is that children seem initially unable to carry out the rather rare VP-fronting operation. Thus, among roughly 40 children at the age of 4 years, most repeated:

(32a) Fussball zu spielen macht Spass
       [to play football makes fun]

as:

(32b) "Fussball zu spielen, <u>das</u> macht Spass"
       [to play football, <u>that</u> makes fun]

adding a resumptive pronoun with left-dislocation. By contrast, without <u>zu</u> the children rarely inserted <u>das</u>:

(32c) "Fussball spielen macht Spass"

---

[15]In fact, it seems very likely to me that OVS is acquired before 4 years, but the argument remains the same: the rarer elements specifying the left hand variable will not be added all at once, as the Topic example below illustrates.

[football-playing makes fun]

Why? Because the compound constitutes a typical DP while a fronted VP evidently does not fall under a DP label, therefore a DP equivalent (<u>das</u>) must be added, converting the structure into Left-dislocation. Adults allow V2 after any fronted XP, but only when the XP abstraction has been fully projected.

These data indicate with some subtlety the precision with which the V2 rule that children use must be represented. They had not yet fully extended the representation of V2 to have a lefthand XP. They must still be assembling particular local environments, not yet collapsed into a single rule.

Data of this kind is what we need to see at what formal level the rule is being written in the child's grammar. The exact steps in the abstraction process would be good to know because we could then determine if they are repeated in L2. For instance, an English speaker who knows that VP-fronting can occur in English might quickly generalize V2 from Subject, then locative, then Object, to full V2 with any XP in German. If the English speaker learning German passed through the same stages, he might reject V2 with VP-fronting just as a child does.

A real possibility is that while children gradually proceed to add lefthand environments for V2, L2-speakers acquire OVS, then hear VP-fronting and essentially jump to the full XP-V-YP abstraction. This is potentially a very important L2 variation in the acquisition path. It would suggest a topdown bias that might have pedagogical implications. If true, it might correspond to L2 pitfalls as well: domains where too broad a generalization is introduced.

The claim that V2 is not full-blown instantly is evident in the fact that Topicalization and V3 are a known problem for V2 speakers.

TOPICALIZATION, INTERFACES AND V3

Our argument is, once again, that UG supplies Multiple Grammars and that critical rules are stated in Minimal terms which invites overgenerating abstractions.

Such a system seems extremely unconstrained. Therefore, we claim that constraints must be present, but from a much different source—not limitations on Feature-bundles in the syntax, nor conditions of lexical representations (e.g., verb particles). Instead we will seek Interface conditions —a view advanced by much modern research. We argue that these constraints may be language-specific. The ultimate question is not whether Interface conditions apply, but how we can state them with a precision that produces exact L2 representations.

Topicalization is an appropriate test case for this interface question. It has been observed that foreigners learning German have difficulty with V2 when Topicalization occurs, and in parallel German speakers have difficulty blocking V2 with Topics (that is producing V3) from a German perspective (see Meisel, 2011 for a literature overview):

(33) a. V3: John I like
      b. V2: John like I

It is easy to see this issue in purely syntactic terms of whether V2 has applied or simply V-to-T.

Rankin provides several citations for the unsurprising claim that V3 is an L2 challenge:

"Meisel (2011, p. 132) points out that "[ungrammatical V3 order] represents a particularly persistent pattern in the speech of L2 learners of German." This is supported by research on L1 English-L2 German by Beck (1998), who finds that learners have persistent problems with the V2 pattern. The influence of L1 word order is thus known to persist at higher proficiency levels."

Here we shall argue that much more than the level of abstraction for XP in a V2 rule is involved.

Information Structure and Strict Interfaces

The status of Topicalization in a grammar has deep roots in issues of Contrast, Focus, and Exhaustivity, which link to current work on Information structure that remains both intricate, unresolved, and sensitive to language-particular variation. We discuss several offshoots before we turn to Topicalization.

Our first goal here is to pose questions which respect the potential role of this interface. Sorace (2011) suggests that L2 variation is vulnerable to indeterminacy at the interfaces and we consider this a valuable and plausible hypothesis. The hypothesis, nonetheless, should be constrained by developing a larger tapestry.

In Roeper (2014) an argument is advanced for Strict Interfaces which I argue here should be present as a backdrop to any claims about Interface variability. The claim is that certain fundamental interfaces must, quite obviously, be presupposed as universal: we assume that phonology links to syntax and syntax links to semantics. That is, humans are not parrots who can master only phonology. We suggest that at a more refined level, UG has the following constraint:

(34) UG obeys Minimal Modular Contact

That is, in the ideal case, two modules have one point of contact through which information flows, which vastly restricts the set of possible syntax-semantics-pragmatics mappings that a theory of interfaces can automatically imply. Consider the notion of Agency. It can be found in the projection of verbs in morphology (-er), projection of roles onto syntax (subject position), and via implicit arguments in the passive.

However, each of these dimensions is mediated by the verb:

(35) Verb maps AGENT onto: Subject position
                                             Implicit Agent
                                             -er

Therefore, -er does not carry Agency by itself, but only if licensed by a verb, which also projects Subject-Agents and implicit agents.

The point becomes clear when one considers child examples like:

(36) "I'll be the listener and you be the storier" (Maria Roeper)

which a child said, but no longer says. Why is such a handy and natural noun (storier) dropped? While –er could be identified with AGENT and therefore attach to nouns, UG demands that the AGENT-role must be linked to specific verbs and projected from the Verb—which has a single point of insertion (hence contact) in the sentence, from which it projects onto

the morphology (-er), syntax (subject), and semantics (implicit argument structure). A child will drop storier when the verbal interface is built and the constraint obeyed[16]. This property of verbs as the Contact point between a dimension of semantics (Argument structure) and syntax is presumably universal.

Imperatives

Likewise there is a natural interface for imperatives between syntax (delete you), semantics (imperative force), Pragmatics (visual situation) and stress intonation (emphatic verb). Very young children understand:

(37) "don't" [applies to child's action]

It seems natural to assume that the imperative interface is largely innate. And the connection between Contrast and Stress and the semantic projection of sets could be innate, although delayed until children have the world knowledge to project appropriate sets. Thus, the capacity to substitute for the stressed word producing different sets could easily be innate:

(38) a. Don't throw BIG STONES.
         b. Don't THROW big stones

creates separate verb and noun sets. Neither the intonation pattern nor the appropriate sets are UG-fixed, but the interface among them could be. Therefore, the "variable" interface itself might be quite small, though significant nonetheless. We shall try to further reduce the variability by claiming that it is not random but reflects only grammatical choices.

The Discourse Option

So where does Discourse reference belong in this realm? Information structure, primarily in terms of Givenness, has been prominently alluded to in Scandinavian studies Eitler and Westergaard (2014)[17]. We argue below that this approach needs to be enriched to include a full description of the Interface and factors like Exhaustivity and Contrast.

Work in L1 has suggested, from several perspectives (Rizzi, 2000; Yang, 2002; Hyams, 2011; and others) that if children's grammar begins in a way dominated by context and discourse, then they should allow Topic deletion (which we discuss in obviously rather simplified terms (see Sigurdsson, 2011 for some discussion). One should see if this extends to L2 speakers as well. In fact, most English speakers are not uncomfortable with discourses where subjects are deleted because they are identical to Topics (Perez et al., 2007):

(39) a. X: have you seen John anywhere?
         b. Y: __went outside a few minutes ago.

---

[16]Note that nouns can take –er: New Yorker, Detroiter while other nouns take –or (donor), and many idioms arise sinker (see Roeper (2014) for discussion), but their readings are idiosyncratic. See Bauke and Roeper (2011) for discussion of where compositionality in the lexicon can be found systematically.

[17]Eitler and Westergaard (2014) argue that the choice of V2 vs. non-V2 in OE/ME was dependent on information structure: The word order XSV (non-V2) was preferred if the subject was informationally given (often a pronoun) and XVS (V2) if the subject was focused or new information (often a full DP).

while closely related forms would seem faulty:

      c. Y: * __is outside.

Why? Perhaps because "outside" by itself is available. "Outside" answers a hidden question-under-discussion "is he inside "anywhere?" Maybe, re-projection of a new Question-under-discussion is preferable to an empty subject. Thus, the application of this Topic-drop principle, not a core part of English, shows subtle variation. Knowledge of such variation we might not expect of an L2 speaker. Would a German L1-English L2 speaker judge (38 b.c) the same way? Or would "__ist draussen" (is outside) be just as good for her? For me, an English L1-German L2 speaker, no clear judgment is available, but I would guess that it is more acceptable[18].

Such issues interact with V2. Consider the environments where exhaustivity arises (see Schulz et al., 2015) for a refined discussion. Although it is not clearly universal (French and Mallayalam are reported to be exceptions), clefts imply exhaustivity in English (See Kiss, 1998; Heizmann, 2012) which children do not initially grasp:

(40) a. it was the dog that ate the cheese

(39a) implies that no one else ate the cheese. It has been suggested that Topicalization also carries exhausitivity, either via a real Operator or as a presupposition at another level:

(41) stones, John picked up.

means he picked up nothing else. However, the sentence:

(42) John picked up stones.

has a weak implicature that nothing else was picked up (Kratzer, 2009), but it certainly does not carry this as a part of its truth value.

Does this hold for grammars where Topic is more generally applied such as German? We may not have a definitive answer at the moment. Nevertheless, what should we expect of an L2 speaker coming from a Topic-dominant L1? Would both of these be grammatical without an exhaustivity expectation:

(43) Dogs, Jim likes
      Dogs likes Jim

We might expect that the L2 speaker will in fact generate both options, but use the potential Information Structure difference as the basis for a choice. What could that difference be?

Let us make two simplified assumptions (whose simplicity might correspond to L2 assumptions), based on the discussion above, and then build an artificial interface which an L2 speaker

might also build. We develop this idea for demonstration purposes only, not as a claim about these language families:

(44) Non-Topic oriented language:
                Topicalization is: a) contrastive
                             b) exhaustive
      Topic-oriented language:
                No contrastivity: no exhaustivity
                      Syntactic V2

Suppose an English speaker acquiring German hears:

(45) Hunde mag Hanns [dogs likes John]

but makes no special assumptions. Then he wishes to express Contrast or Exhaustivity via Topicalization. He might then in German utilize an English device to indicate exhaustivity, saying incorrectly:

(46) Hanns mag Tiere nur selten, aber
        Hunde, Fritz mag.
        [Hanns likes animals only rarely, but dogs Fritz likes]

On other occasions, V2 could arise where this implication was immaterial as in (44). Thus, apparent variability at the interface could be resolved into distinct choices available to the L2 speaker applying MG, but not the monolingual speaker.

The reader can see how this toy scenario works. We do not have to assume that there is pure indecision leading to variability, but rather, at a subtler level, we apply MG theory, via an available UG interface option, which creates two options. An L1/L2 speaker uses both depending upon the interface circumstances, thus never using a "variable" grammar.

An interesting challenge here would be to design experimental scenarios that might elicit these distinctions.

(47) [scenario: John catches fish but also a turtle]
      a) "Fish caught John"
      b) "Fish John caught"

Now if only (46b) is exhaustive, then the L2 speaker might say (46b) in German in order to capture the exhaustivity. On other occasions where only emphasis is sought, we would find (46a).

In other situations, where the meaning is not grammatically captured, then it must be otherwise unreliably inferred. Kratzer (2009) suggests that there is a hidden equivalent of only at the pragmatic level. This approach to Interfaces claims that what looks like variability might be an effort to impose greater semantic exactitude through L2.

We can now enlarge our realm of possibilities to include this strong claim, which is useful in framing the acquisition problem even if it proves questionable:

(48) Languages may have Unique Interfaces

That is, the combination of syntax, semantics, and pragmatics might involve an implication in a particular language that is unavailable directly in other grammars (although surely communicable by more indirect means).

Suppose for instance there is an Honorific in a language and a Topic rule, such that we combine Exhaustivity with an implication that the Honored person must be present. Then we

---

[18]Hyams (2011, p. 42) recounts a number of L1 studies that show Discourse sensitivity: "Allen (2000) shows the argument omission vs. overt expression can be significantly predicted by the degree of "informativeness" of an argument (as measured by several variables including newness, contrast, absence, differentiation in context and person). Serratrice and Sorace (2003), using the same principles introduced by Clancy and Allen, also find significant discourse/pragmatic effects in the distribution of overt vs. null subject in six Italian-speaking children (ages of 1:8 and 3:3), reflected the distribution of the adult language. Serratrice and Sorace are explicit in assuming that the pragmatic principles operate within the boundaries imposed by the grammar, in this case a pro-drop grammar."

arrive at the meaning: only one such person is present now. This approach supports the intuition that while anything can be said in any language, some meanings might be grammatically expressible via grammar in one language that must be explicitly asserted in another. We can conclude that if there are unique Interfaces, then it is exactly L2 and Heritage language research which may be able to isolate them.

Expletives and MG

While one might suppose that Topicalization rules out V2-like inversion altogether in English, this is not the case. Consider this contrast:

(49) there are three bananas.
  a. Only two of them is it good to eat
  b.*Only two of them, it is good to eat.

Expletives do not seem to allow Topicalization in English without inversion. What would the German L2-speaker of English think? Here we might imagine exactly that the syntactic availability of both forms could mislead the advanced speaker who restricted V2 for Topic, giving V3 in English, into saying or accepting V3 or (b) when that would be a mistake. This would be an example of an L2 speaker applying an overgeneral V3 rule that allowed expletive to follow a Topic.

MISSING SUBJECTS AND OBJECTS

Our focus on MG and modular compatibility has focused on V2, but we will briefly note that there are two other domains where an analysis in another grammar does not disturb other modules: empty subjects and objects.

Perez et al. (2007) reports that children will misanalyse empty generic objects as discourse-linked, which is grammatical in Spanish and Portuguese. Consider this situation:

(50) Scenario: mother is cooking eggs.
      Child: look Mom, I caught a fish.
      Ask Subject: Is the Mom cooking___?

Spanish children and even English-speaking children initially say:

"no" because the Mom is not cooking fish, filling in the object of cook with a contextually salient object. Spanish speakers of L2 say that they must actually suppress this reading [Luiz Amaral (pc)] in order to favor a generic object [cook (something)], to which the answer is "yes" (since she is cooking eggs). Note that this case satisfies Compatibility because there is no misanalysis in another module.

Likewise missing subjects can be used by an L2 speaker without disturbing another module:

(51) Where's John => "___is singing"

and therefore is predictable in this theory.

Modular Incompatibility

What happens where there is incompability with another module. We can generate a prediction[19]. Schouwenaars et al.

---

[19]There may well be relevant data available on this question of which I am not aware.

(2014) report that Dutch 5 year olds will overapply SVO analysis to object-fronted sentences even when the subject-verb agreement should force an OVS reading:

(52)   who wash-plural the dancers
      => who are the dancers washing t

This result might be found among L2-speakers or via eye-tracking which would indicate that an SVO analysis operates at a superficial level and then undergoes revision as new modules are added. This interaction among modules might well be most visible via research on L2. It could lead to quite subtle degrees of acceptability.

Minimalism and Abstraction

A general consequence of Minimalism is that rules are stateable at a very abstract level. One can, for instance, build structures with decisions about Labels left partly open. This creates extra L2 room for uncertainty[20].

Where else can we find evidence of the abstract level of a rule? Here is a case one might subject to greater scrutiny. In English we find many speakers (including me) who say things like (from a Google for "could have I/you"):

(53) "How could have I passed the exam"
      "How could have you done this to me?"
      "how could have you used your powers for evil?"

instead of:

(54) How could I have passed the exam

Will an L2 speaker allow both in comprehension or production? The answer most probably lies in whether the grammar represents inversion with an Aux-Head or an unspecified AuxP:

(55) NP AUXP/AUX V

This is an empirical question, but if the approach advocated here is correct, then speakers should aim for more abstract representations rather than less abstract ones. Therefore, the AUXP inversion will probably not be rejected so easily by L2 speakers, even if not used.

In a sense we can characterize the L2 acquisition path as topdown rather than bottom up. If the child builds up a very narrow range of possible environments initially and finally generalizes to a full range of invertable Auxiliaries, the L2 speaker might seek to build the most abstract form as quickly as possible as an instance of representational economy.

Dialects and Compatibility

Is it impossible to write features of one grammar into another? Green and Roeper (2007) argued that one way to define a dialect is in terms of Tree-compatibility. Green has argued that

---

[20]Here is typical anecdotal case of L2 leaving a wh-node without a feature. A foreign student once asked whether whose in English must be a person because it has who within it which by itself does have this restriction. In fact, of course, it is not required, but the evidence for this may not arise everyday and therefore one would seem entitled to continue the assumption that the morpheme who had a [+person] restriction.

there is an Aspectual node in African-American English. It can be added between IP and VP in Mainstream American English without disturbing the tree, but with discernible consequences:

(56) a. He be playing baseball, don't he.
      b. *He be playing baseball, ben't he/isn't he.

Here we find the tag-question indicates that the habitual be belongs to an extension of VP, not IP, therefore requiring do-insertion, just like a Main Verb. Therefore, it must belong to the Verb-projection, not IP:

(57) He plays baseball, doesn't he.

Non-AAE speakers understand and sometimes use Habitual BE, but fail to form do-tags, suggesting that they assimilate it to IP and not VP. In any case, the dialect speaker who also controls the Mainstream form will need to have diacritics to indicate social factors that dictate whether the extra node should be allowed.

Variational Learning, Feature-Reassembly, Full T/Full A

In a sense, the MG approach is a methodological proposal orthogonal to, not in opposition to, current theories. The essential proposal is simply to formulate the grammars of L2 with sufficient technical precision that they predict what is ungrammatical in the manner of L1 research. To capture the formal "variability" one should state as rules or grammars the options selected. Of course, whenever formal variability arises, it invites myriad social and pragmatic factors to participate, producing the surface variability of sociolinguistic "optionality," which—if we understood them fully—may or may not be represented as Features in the formal rules.

Variational Learning (VL)

The VL approach has MG as a prerequisite[21]. Yang (2002) argued that each side of a parameter—both of which must be present—is linked to a probability which is increased or decreased by further evidence. It could not exist if one attempted to represent the facts within one grammar with complex exceptions. The unchosen—or non-productive side remains in an available grammar.

Yang argues that the weight on one side of a parameter over another is increased or decreased in terms of input experience. An interesting question to ask here in this light is whether one is responding primarily to types or tokens.

Consider the pro-drop parameter which is arguably triggered by sufficiently frequent exposure to one type, there-insertion. When a child hears enough examples of it[22], then English

is represented as -Pro-drop. However, many, many examples, like:

(58) a. seems nice
      b. looks good.

exist so that the +pro-drop parameter seems to survive linked to specific verbs, no matter how frequent they are.

In the case of V2, as we have discussed, it is the types of constructions which can occupy the lefthand XP position which seem to be critical to the eventual productivity of the expression.

Nevertheless, the English speaker also operates with verb classes so that the fairly large class of speaking verbs uniformly permits it:

(59) "nothing" roared/muttered/sighed/moaned Bill

And the verb be is extremely productive and compatible with V2. We say:

(60) a. How is it

and not:

      b. *How do it be.

A brief search in CHILDES revealed 6 children who appear to generalize this to the category of equatives and say:

      c. "what means that"

or a period before it is eliminated by hearing "what does that mean." The large number of be sentence tokens, however, does not trigger generalized V2 as in German. Therefore, the type/token difference is important. We do not yet know how to conceive of the balance between them in order to determine productivity. Is it the many types or many tokens which etch a rule into a grammar?

Feature-Reassembly

Another approach is to reduce all variation to features which can then be variously valued as proposed by Feature-Reassembly (Lardiere, 2009) who provides insightful efforts to apply modern linguistic distinctions to L2. Often it is not exactly clear where the weight should fall: feature choice, uninterpretability, feature-assembly, morphology, or meaning variation. Lardiere (2009) shows quite well how the theories of parameters and micro-parameters overgenerate, providing an insurmountable range of options, and do not make precise predictions. We agree with her apt summary:

"Parameter-setting, however, has never coped very well with the issue of variability, which is often a persistent hallmark of second language development. (By "variability," I mean here the variable omission, underspecification, overreliance on default forms, and/or apparent optionality vs. obligatoriness of the morphophonological expression of grammatical properties.) As van Kemenade and Nigel (1997) point out, since parameter settings are typically all-or-nothing phenomena, the resetting of a parameter should represent an "abrupt change" in a speaker's language (p. 4). The persistence of observed variability in the acquisition data is thus not predicted, insofar as the presence

---

[21]See Kroch and Taylor (1996) for connections to historical grammar and the earliest Multiple Grammar proposals for generative grammar. As one reviewer suggests, the approach should naturally apply to intermediate stages in the history of grammar. See Yang and Roeper (2011) for a broad background.

[22]See Hyams (2011) for a current overview which also articulates other dimensions besides a single trigger. While other factors may support or be a pre-requisite, it is not clear that a few central triggers are not the basic pivot around which a parameter is set. See Holmberg (2010) for a sophisticated presentation along these lines.

or absence of some grammatical property should be tied to the learner's having set the plus or minus value of a particular parameter."

This critique leads us precisely to think that we need a conception of L2 and language variation that is pitched at principles expressed at the macro-level: Head direction, wh-movement, LF variation. Once these choices are formulated as independent grammars, which are all present in everyone's UG, variability follows naturally (even if conditioning social, phonological, and pragmatic factors are difficult to state).

What is being proposed here is not in opposition to these potentially useful notions from Feature-reassembly which appropriately argues that feature addition and subtraction are insufficient. Rather, once again, MG offers a different methodological approach. It suggests that L2 research, whatever mechanism and formalism is involved, should proceed from exact formulations that arrive at predictions of acceptable or unacceptable grammaticality for an L2 speaker. This is how generative grammar began: very simple, now almost quaint formalisms in early work by Chomsky, but a steady refinement of them with predicted and rejected instances of acceptability/grammaticality. Without a sharp edge of this kind, I believe it will be difficult to build the kind of theory of L2 acquisition that most researchers would like to see.

Thus, to capture variation at that level, particularly that which reflects both L1 and L2, we need to write out two independent grammars and claim that they are both active. In order to do that, one needs particularly abstract representations—exactly of the sort we have been discussing. Access to the abstraction as a starting point is critical.

If one can write the grammar with abstract notions like Maximal Projection (XP), then one can begin to state the variations as we have done above:

(61)   CP
           /   \
         Spec   C
          | | \
         XP V_1 YP
                      /   \
                    Y [VP [V $trace_1$]]   [Y = any material]

where VP can allow variation under a modified Head Constraint to include V-particle, or Aux-Head and AUX-complex.

We need a perspective more abstract than Feature-reassembly to capture this. Consider the prediction Lardiere (2009) makes that a Chinese speaker confronted with "I bought fruit" will give a Real question answer rather than an Echo-question:

(62)   a. A. "I bought fruit"
          B. "you bought what" = echo-question =>
          A. "I bought fruit"
       b. A. "I bought fruit"
          B. "you bought what"
              = Real question, seeking specification
              [= "what did you buy"] ⇒
          A. "I bought bananas."

(b) asks for greater subdivision of given knowledge, a subset answer. Why this prediction? It is clear from her results that both options are available, so we need to state them both as alternatives.

Full Access/Full Transfer

The original proposal of FA/FT by Schwartz and Sprouse (1996) launched a tremendous amount of detailed work seeking every hint of cross-linguistic effects in the L2 process. This large net is a natural first step and very valuable. However, the concept seems to presuppose that one inserts part of a grammar into another with a great deal of minute adjustments then following and perhaps a great deal of L1 baggage that does not fit. FA/FT does not have a natural way to capture the dual analysis of particular simple sentences like those cited from Rankin: who woke John up. And it does not have a metric to describe the diverse impact of different levels of grammar. Perhaps one should see MG as beginning to carve out a space for such metrics which could reflect Interface boundaries (as Sorace suggests), although we regard much of the interface domain to be universally determined and precisely where little variation occurs.

To appreciate one case where Transfer is examined, consider Özçelik (2009) who looks at Inverse Scope in Turkish:

(63)   a. Donald didn't find two guys = inverse two guys>not
                [=there are two guys Donald did not find]
          surface: not> two guys
                [Donald did not find (any) two guys]

The author comments:

"intermediate English L2ers should behave noticeably worse than advanced English L2ers due to the initial transfer of the Turkish setting, as well as the ongoing acquisition of the L2 setting. However, our intermediate English L2ers did not particularly do bad enough to be qualified as "transferring from the L1.""

Inverse scope is a major option in the organization of LF and therefore we would expect it to be among the abstract rules that is available as a separate entity from UG with minimal triggering required. If we can assume, therefore that if they have any evidence that invites Inverse Scope, then that grammar will make Inverse Scope available and it can apply. It is not a question of whether it came from Turkish or whether it is transferred to English, but simply whether evidence has arisen to instantiate that important UG option. Once present we would expect it to remain as a comprehension option even if speakers were able to avoid it in production. The Comprehension/Production distinction is particularly important for L2 (see Amaral and Roeper, 2014).

Typological Primacy Model and Bottleneck Theory

Rothman's (2013) presentation of the Typological Primacy Model shares much with our approach, in particular the desire to make strong predictions and to argue that what is transferrable depends upon the grammatical status of constructions. Generating predictions strong enough to be proven "wrong" is the traditional path to refinement in the history of generative grammar.

While the TPM puts an emphasis on the intertwined nature of syntax, morphology, and phonology, we argue that it is exactly the extent to which properties of a given module can be cast in an abstract independent form—be un-entwined—that will dictate their transferability. In that light, as Rothman points out, LF transfer works cross-linguistically. He also cites Özçelik (2009): "[who] argues explicitly for and shows convincing evidence of overall typological and not property level structural transfer in line with the TPM, showing that Uzbek–Russian bilinguals of L3 Turkish transfer scopal properties of Uzbek, a Turkic language like Turkish, despite the fact that Uzbek works differently and Russian and Turkish are identical in this regard."

Likewise his notion of "degree of similarity" refers ultimately to the degree to which different modules are intertwined, so that the less other modules are involved, the more likely a simple, transferrable rule is possible. The TPM refers to this notion as "non-facilitative" which is the same prediction our account makes. The challenge, as always, is to define sharp representational options for what is claimed.

We can extend the LF example further in terms of interaction with case-marking. If case-marking is universal, but can be abstract and show no morphological effect, as is generally the case in English, then it will not interfere with LF formation, therefore Transfer should occur.

(64) someone loves everyone => LF [everyone > someone]

But note that the theory could be shown to be wrong if one language shows no case-Marking while another language marks quantifiers with case, like German:

(65) Jemand liebt jeden [object-case marking] (someone loves everyone)

If then transfer to or from German with case-marking of Scope inversion is more difficult than transfer to Chinese without case-marking, then it would show that LF does not have modular independence. And if a language marks both nominative and accusative on quantifiers, then we might predict even less LF transfer. If so, then we would have evidence of interference, presumably blocking use of LF scope inversion. But again, if LF has a case-independent representation, then it should transfer in all languages equally. These are, clearly, easily approachable empirical questions.

Moreover, this example may be a domain where we can fulfill the promise that cross-linguistic comparative work can further articulate UG. It is safe to say that there is a common intuition that LF movement has nothing to do with case-marking: we do not have to move invisible case-marking when we covertly move a quantifier. If there is no contrast between LF in case-marked and non-case-marked languages for transfer, then it is direct evidence for this intuition, which should ultimately be stated in a fully-articulated representation of UG.

Consider now Bottleneck theory. Slabokova (2014) argues that the involvement of Functional Categories (FC) proves difficult to transfer across languages. Again the generalization implicitly refers to the fact that FCs (e.g., CP) can engage other modules, like wh-movement. If we have a sentence like:

(66) Whom did you talk to__

We have not only the projection of CP, but a Question-Probe feature which causes wh-movement to occur, but only after case-marking has applied. The fact that several modules are involved is doubtless related to the fact that case-marking is weakening in this construction and allowing who did you talk to_ for many speakers. On the other hand, direct lexical expressions of FC's (like complementizers that or to) may show minimal transfer inhibition or delay in acquisition. Once again we argue that it is the interaction of several modules that may block easy application of one grammar inside another as it is formulated in MG. Such interactions may be very common in FC's, but it may not be the concept of FC itself which is the source of difficulty.

CONCLUSION

Let us summarize our approach. The MG theory is, in a basic sense, an inevitable consequence of the abstract nature of modern minimalism. It means that via abstraction one can state common rules across many grammars. This is a more powerful UG claim than the traditional view that the building blocks of all grammars are identical.

Another emphasis in this essay is that many of the MG options remain at an abstract level and are constrained by unique interface restrictions rather than restrictions stated on the rule itself. Our goal has been to propose that if we articulate full MG options that include fixed Interface representations, avoidance of other modules that complicate the application of rules, we will have a method to generate more precise acceptability/grammaticality judgments from L2 speakers. In this approach, the notion of Transfer is supplanted by explicit presence of two analyses whose status can be experimentally explored.

It follows naturally that if we allow ourselves more abstract representations, then those representations lend themselves to the idea that a rule can apply across grammars, or that alternative rules (V=>Tense, V=> Comp) are jointly available for both monolingual and bilingual speakers. These questions can be approached applying detailed experimental apparatus, which we have presented here in a speculative manner. Altogether, it should be clear that a whole phalanx of predictions arise from the MG account.

This leads to what might seem like a paradoxical result. Although one might say that the presence of two grammars should make analysis more obscure and ambiguous, the argument here is that it is precisely this assumption, used in L2 research, which can isolate fundamental properties of grammar where monolingual analysis permits too many alternatives to make a decisive choice. If successful, then research on multi-lingualism holds the promise of theoretical insights unobtainable anywhere else.

REFERENCES

Adger, D., and Smith, J. (2005). "Variation and the minimalist program," in *Syntax and Variation: Reconciling the Biological and the Social*, eds C. Leonie and P. C. Karen (Amsterdam: John Benjamins), 149–178.

Alexiadou, A., and Anagnostopoulou, E. (2007). "The subject-*in-Situ* generalization revisited." in *Hans-Martin Gärtner and Uli Sauerland, Proceedings of the Workshop on Interfaces + Recursion =Language?* (Berlin: De Gruyter Mouton), 31–60.

Allen, S. (2000). A discourse-pragmatic explanation for argument representation in child Inuktitut. *Linguistics* 38, 483–521. doi: 10.1515/ling.38.3.483

Amaral, L., and Roeper, T. (2014). *Multiple Grammars and Second Language Representation Second Language Research, Vol. 30* (Sage Journals).

Bauke, L., and Roeper, T. (2011). "How phase-based interpretations dictate the typology of nominalizations," in *Grammar and Discourse. From Sentence Types to Lexical Categories*, eds G. Grewendorf and T. E. Zimmermann (Berlin: De Gruyter Mouton), 289–320.

Bruening, B. (2013). *Quotatives*. Berlin: Univerity of Ulm.

Chomsky, N. (1957). *Syntactic Structures Mouton.* Berlin: De Gruyter Mouton.

Chomsky, N. (1995). *The Minimalist Program.* Cambridge, MA: MIT Press.

Chomsky, N. (2013). Problems of projection. *Lingua* 130, 33–49. doi: 10.1016/j.lingua.2012.12.003

Collins, C., and Branigan, P. (1997). Quotative inversion. *Nat. Lang. Linguist. Theory* 15, 1–41.

Curtiss, S., Froimkin, V., Krashen, S., Rigler, D., and Rigler M. (1974). The linguistic development of Genie. *Language* 50.3, 528–554.

Eitler, T., and Westergaard, M. (2014). "Word order variation in late Middle English: the effect of information structure and audience design," in *Information Structure and Syntactic Change in Germanic and Romance Languages,* eds B. Kristin and E. Kristine (Amsterdam: John Benjamins), 203–231.

Green, L., and Roeper, T. (2007). "Node labels and features: stable and unstable dialects and variation in acquisition yearbook of variation," in *Yearbook of Morphology*, ed G. Booij (Springer), 1–26.

Heizmann, T. (2012). *Exhaustivity in Questions and Clefts and the Quantifier Connectioin: A Study in English and German*. Amherst, MA: UMass Dissertation.

Holmberg, A. (2010). Parameters in minimalist theory: the case of Scandinavian. *Theor. Linguist.* 36, 1–48. doi: 10.1515/thli.2010.001

Hyams, N. (2011). "Missing subjects in child language," in *Handbook of Generative Approaches to Language Acquisition, Studies in Theoretical Psycholinguistics*, eds J. de Villiers and T. Roeper (New York, NY: Springer), 13–52.

Keyser, S. J., and Roeper, T. (1992). Re: the abstract clitic hypothesis. *Linguist. Inquiry* 23, 89–125.

Kiss, K. (1998). Identificational focus versus informational focus. *Language* 74, 245–273.

Kratzer, A. (2009). "Pragmatic strengthening for free. Plenary talk," in 31. *Annual Meeting of the German Linguistic Society (DGfS)* (Amherst, MA).

Kroch, A., and Taylor, A. (1996). "Verb movement in old and middle English: dialect variation and language contact," in *Parameters of Morphological Change,* eds A. van Kemenade and N. Vincent (Cambirdge, MA: Cambridge University Press), 297–335.

Lardiere, D. (2009). Some thoughts on a contrastive analysis of features in second language acquisition. *Sec. Lang. Res.* 25.2, 173–227. doi: 10.1177/0267658308100283

Lohndal, T. (2013). Generative grammar and language mixing. *Theor. Linguistics* 39, 215–224. doi: 10.1515/tl-2013-0013

Meisel, J. (2011). First *and Second Language Acquisition.* Boston, MA: Cambridge University Press.

Müller, N. (ed.). (2003). *(In)Vulnerable Domains in Multilingualism.* Amsterdam: John Benjamins.

Özçelik, O. (2009). "L2 acquisition of scope: testing the full transfer full access hypothesis," *Proceedings of the 10th Generative Approaches to Second Language Acquisition Conference (GASLA 2009)*, ed B. Melissa et al. (Somerville, MA: Cascadilla Proceedings Project], 168–179.

Perez, A., Pirvulescu, M., and Roberge, Y. (2007). *Null Objects in Child Language: Syntax and the Lexicon Lingua*. Boston, MA: Elsevier.

Rankin, T. (2015). *Variational Learning in L2: The Transfer of L1 Syntax and Parsing Strategies in the Interpretation of wh-questions by L1 German learners of L2 English*. Vienna University of Economics and Business.

Rankin, T. (2013). *Competing Grammars in the Comprehension of Questions and Relative Clauses in L1 German—L2 English*. Ms. Vienna University of Economics and Business.

Rizzi, L. (2000). "Remarks on early null subjects," in *The Acquisition of Syntax*, eds M. Friedemann and L. Rizzi (Harlow: Longman), 269–292.

Roeper, T. (1973). *Approaches to the Acquisition of German.* Ph.D. Dissertation, Harvard University, Cambridge, MA.

Roeper, T. (1999). *Universal Bilingualism. Bilingualism, Language and Cognition.* Cambridge, UK: Cambridge University Press.

Roeper, T. (2011). "Strict interfaces and three kinds of multiple grammars," in *The Development of Grammar. Language Acquisition and Diachronic Change*, eds E. Rinkle and T. Kupisch (Amsterdam/Philadelphia: John Benjamin), 205–228.

Roeper, T. (2014). Strict interface principles and the acquisition engine: from unlabeled to labeled and minimal modular contact. *Lang. Sci.* 46, 115–132. doi: 10.1016/j.langsci.2014.05.003

Rothman, J. (2013). Linguistic and cognitive motivations for the typological primacy model (tpm) of third language (l3) transfer: timing of acquisition and proficiency considered. *Bilingualism* 18, 179–190. doi: 10.1017/S136672891300059X

Schouwenaars, A., van Hout, A., and Hendriks, P. (2014). "Word order overrules number agreement: Dutch children's interpretation and production of which-questions," in *Selected Proceedings of the 5th Conference on Generative Approaches to Language Acquisition North America (GALANA 2012)*, eds C-Y. Chu, C. E. Coughlin, B. Lopez Prego, U. Minai, and A. Tremblay (Somerville, MA: Cascadilla Proceedings Project), 60–71.

Schulz, P. A., Müller, A., and Lange, C. (2015). Exhaustivity in wh-questions and only-focus: same or different? *GALA Presentation.*

Schwartz, B., and Sprouse, R. (1996). L2 cognitive states and the Full Transfer/Full Access model. *Sec. Lang. Res.* 12, 40–72. doi: 10.1177/026765839601200103

Serratrice, L., and Sorace, A. (2003). "Overt and null subjects in monolingual and Bilingual Italian Acquisition," in *Proceedings of the 27th Annual BUCLD,* Vol. 2 (Somerville, MA: Cascadilla Press), 739–750.

Sigurdsson, H. (2011). Conditions on argument drop linguistic. *Inquiry* 42 267–304. doi: 10.1162/LING_a_00042

Slabokova, R. (2014). "The bottleneck of second language acquistion foreign language teaching and research monthly, Vol. 46.4," in *Fourth International Symposium on Chinese/Applied Linguistics* (Iowa City, IA).

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. *Linguist. Approach. Bilingualism* 1, 1–33. doi: 10.1075/lab.1.1.01sor

Vainikka, A., and Young-Scholten, M. (1996). Gradual development of L2 phrase structure. *Sec. Lang. Res.* 12, 7–39.

van Kemenade, A., and Nigel, V. (1997). *Parameters and Morphosyntactic Change.* Cambridge University Press.

Westergaard, M. (2003). "Unlearning V2: transfer, Markedness and the importance of input cues in the acquisition of word order in English by Norwegian children," in *EUROSLA Yearbook*, Vol. 3, eds S. H. Foster-Cohen and S. Pekarek Doehler (Amsterdam: John Benjamins), 77–101.

Wexler, K. (1998). Very early parameter setting and the unique checking constraint: a new explanation of the optional infinitive stage. *Lingua* 106, 23–79.

Wexler, K. (2011). "Grammatical computation in the optional infinitive stage," in *Handbook of Generative Approaches to Language Acquisition*, eds J. de Villiers and T. Roeper (Drodrecht: Springer), 53–118.

White, L. (2003). *Second Language Acquisition and Universal Grammar.* Cambridge: Cambridge University Press.

Yang, C. (2002). *Knowledge and Learning in Natural Language.* Oxford: Oxford University Press.

Yang, C., and Roeper, T. (2011). "Minimalist approaches to language acquisition" in *Handbook of Minimalism*, ed C. Boeckx (Oxford University Press), 551–574.

frontiers
in Psychology

Neurolinguistic measures of typological effects in multilingual transfer: introducing an ERP methodology

Jason Rothman[1,2]*, José Alemán Bañón[3]* and Jorge González Alonso[4]

[1] School of Psychology and Clinical Language Sciences, University of Reading, Reading, UK, [2] Department of Language and Linguistics, UiT The Arctic University of Norway, Tromso, Norway, [3] Basque Center on Cognition, Brain and Language, Donostia-San Sebastian, Spain, [4] Department of English and German Philology, University of the Basque Country, Vitoria, Spain

This article has two main objectives. First, we offer an introduction to the subfield of generative third language (L3) acquisition. Concerned primarily with modeling initial stages transfer of morphosyntax, one goal of this program is to show how initial stages L3 data make significant contributions toward a better understanding of how the mind represents language and how (cognitive) economy constrains acquisition processes more generally. Our second objective is to argue for and demonstrate how this subfield will benefit from a neuro/psycholinguistic methodological approach, such as event-related potential experiments, to complement the claims currently made on the basis of exclusively behavioral experiments.

Keywords: third language (L3) acquisition, transfer, event-related potentials (ERPs), agreement, artificial language

Introduction

Empirical investigations into adult multilingual acquisition have been done for decades and from a multitude of paradigms (see De Angelis, 2007; Edwards and Dewaele, 2007; Rothman et al., 2013 for review). Prior to the last decade or so, it was not obvious that the study of a third or more languages in adulthood should constitute its own subfield of acquisition research, that is, distinct from the study of a non-native second language (L2). As Edwards and Dewaele (2007, p. 221) state, there is a "growing awareness that trilingualism is not just an extension of bilingualism," meaning that the idea that studying multilingualism simply presents more of the same as bilingualism no longer prevails. It is now definitively clear that there are methodological, cognitive, linguistic, and epistemological reasons why L3 acquisition must be considered independently (see e.g., De Angelis, 2007; cf. de Bot and Jaensch, 2015).

With few exceptions, for example Klein (1995), studies on L3 acquisition of morphosyntax from a formal linguistic perspective did not emerge until the early 2000s. Since then there has been a sharp increase of interest and output of research in adult multilingual acquisition within the generative tradition (see Leung, 2007; Rothman et al., 2011). As pointed out by García Mayo and Rothman (2012), to date much of this work has focused on investigating previous language transfer source(s)[1] under the mindset that doing so is relevant to and provides unique evidence for litigious

---

[1]There is, of course, notable work in generative L3 studies that investigate interlanguage development, regressive transfer effects in development and competence at later stages of acquisition such as, for example, García Mayo et al. (2005), Cabrelli Amaro and Rothman (2010), García Mayo and Villarreal Olaizola (2011), Cabrelli Amaro (2013), Slabakova and García Mayo (2015).

questions that concern all acquisition research. For example, investigating how transfer—influence from previously acquired mental linguistic representations—is constrained in adult multilingualism, where several potential options/sources are available, ultimately contributes to a more fine-grained understanding of underlying linguistic representations and the role of cognitive economy in acquisition processes more generally (see Flynn et al., 2004; Rothman, 2013, 2015 for details).

At present three formal models of L3/Ln morphosyntactic transfer have proved influential in spawning what can now be considered an emerging subfield of generative L3 transfer studies. Not surprisingly given the paradigm in which they are conceived, each of these models is predicated on the notion that multilingual acquisition in adulthood is subject to universal constraints and that transfer in multilingualism is not at all random, but rather is delimited by linguistic and/or cognitive factors. These three models, to be reviewed in greater detail in Section "L3 Models of Morphosyntactic Transfer," are: (i) the *L2 Status Factor* (Bardel and Falk, 2007, 2012; Falk and Bardel, 2011), (ii) the *Cumulative Enhancement Model* (CEM, Flynn et al., 2004; Berkes and Flynn, 2012) and (iii) the *Typological Primacy Model* (TPM, Rothman, 2010, 2011, 2013, 2015). A commonality between them is the shared belief that adult learners are able to acquire new morphosyntactic representations[2] past puberty and that more than strictly speaking linguistic variables (i.e., cognitive considerations) contribute to what ultimately determines selection of transfer and even its timing. Yet, differences in their proposals result in mutually exclusive predictions that render them empirically falsifiable against one another.

Some experimental studies have offered data that are compatible with more than one of these models. This is not surprising since these models do not always offer incompatible predictions depending on the language triad and order of acquisition of the languages under investigation. In the body of this paper, we will introduce and discuss much of the existing empirical data, offering some insights into what we believe they tell us when coupled together. In doing so, we will address the first of two goals of this paper, which is to introduce the reader to this emerging field and the empirical evidence it provides. Since the existing data come exclusively from behavioral methodologies, the second goal of this paper is to show how the methodological remit of generative L3 studies can be expanded to include neurolinguistic methodologies such as event-related potentials (ERPs), as has been done in recent generative L2 work (e.g., Gabriele et al., 2013a; Alemán Bañón et al., 2014). To this end, we will detail how these models make clear predictions that can be tested with an ERP methodology, and articulate

a sample methodology we contend is suitable to test these predictions.

L3 Models of Morphosyntactic Transfer

In the past decade, three generative L3/Ln models of morphosyntactic transfer have been proposed. This section introduces these models, which we propose are testable against one another via processing methodologies, such as ERP.

The L2 Status Factor

As the name suggests, the L2 Status Factor is a model of multilingual transfer which assigns a privileged role to the L2 at the initial stages of L3 acquisition (e.g., Bardel and Falk, 2007; Falk and Bardel, 2011). It is argued that the L1 is not as accessible as the L2 for transfer, presumably because the L2 is represented and stored in a different memory system (declarative memory), relative to the L1 (procedural memory). Falk and Bardel (2011) and Bardel and Falk (2012) adopt a synthesis of Ullman's (2001, 2005) and Paradis' (2004, 2009) Declarative/Procedural (DP) models of bilingualism to offer what they claim to be a neurolinguistic basis for the L2 Status Factor.

The question of why L3 learners would default to suppressing the L1 and rely more heavily on the L2 is of great epistemological importance for the L2 Status Factor. Bardel and Falk (2012) argue that doing so is essentially a byproduct of assumed cognitive similarity between the L2 and the L3. They claim that both the L2 and L3 differ from L1 grammars in terms of the developmental path, the degree of ultimate attainment, and the memory systems they draw from (declarative vs. procedural). In DP models, the grammar of the L1 is sustained by procedural memory (implicit), while declarative or lexical memory (explicit) supports both the L1 lexicon and, at least at the initial stages, the grammar of all late-acquired languages (i.e., L2, L3, Ln). Bardel and Falk (2012) adopt the DP divide of L1 vs. L2 representation and argue that it results in bypassing the L1 as a primary or even possible source of transfer in L3 acquisition.

The data that best support the L2 Status Factor come from Bardel and Falk (2007) and Falk and Bardel (2011). Bardel and Falk (2007) examined placement of negation in two different groups: L1 V2[3]/L2 non-V2 and L1 non-V2/L2 V2, learning either Swedish or Dutch as an L3, both of which are V2 languages. Their data showed that the L1 non-V2/L2 Dutch/German group outperformed the L1 V2/L2 English group in producing post-verbal negation. They maintained that only a privileged role for the L2 is corroborated by the data. Despite compelling evidence that typology was not necessarily a deterministic factor, one must keep in mind that these learners are not beginners and that what we observe could actually be a byproduct of L3 interlanguage

---

[2]Based on the most recent papers by Bardel and Falk (2012), in which they appeal to the so-called declarative/procedural distinction following Paradis (2004), it is no longer completely clear to us that what the L2 Status Factor takes as L2 mental linguistic representation is the same as the CEM and the TPM, the latter of which maintain a clear distinction between learned and acquired knowledge and exclusively focus on the latter type of L2 knowledge. See "L3 Models of Morphosyntactic Transfer" below for further discussion.

[3]V2 refers to verb-second, a distinctive property of Germanic languages (except for English). In V2 languages, the finite verb appears in second position of a declarative main clause, whereby the first position is occupied by a single major constituent that functions as the clause topic. V2 languages do differ with respect to the distribution of the V2 rule, often referred to as micro-parametric variation: while some V2 languages restrict the V2 rule to matrix clauses (e.g., German, Dutch), others have V2 in matrix and subordinate clauses alike (e.g., Swedish, Norwegian).

development itself. That is, it is possible that the pattern would have been distinct if the learners had been tested at an earlier, more appropriate stage in L3 development for the question of transfer source.

Despite plenty of data that clearly show that the L2 is a potential source of L3 transfer, there are less data that unambiguously support the L2 Status Factor's principled claim that it should be the privileged or only source. That is, much of the data showing that the L2 is transferred is not in a position to preclude other variables, such as typological similarity or maximal facilitation, as being the actual deterministic factors for the selection of the L2. The L2 Status factor is clear: despite other variables that might favor the L1 from a typological or facilitative point of view, the L2 should be chosen, precisely due to the neurocognitive reasons detailed above, as cited by Bardel and Falk (2012). Just like showing L1 transfer would only be consistent with absolute transfer under certain methodologies and language pairings, demonstrating L2 transfer might only be consistent with the possibility of L2 transfer as opposed to falsifying alternative explanations. Rothman and Cabrelli Amaro (2010) mention this in their study, which examined properties related to the Null Subject Parameter in L3 French and L3 Italian. Their study could be cited as strong support for the L2 Status Factor insofar as their data show L2 transfer and are thus entirely consistent with the L2 Status Factor's predictions. However, Rothman and Cabrelli Amaro (2010) ultimately concluded that they were unable to differentiate between an L2 Status Factor effect and possible (psycho)typological influences, since the choice of L2 and L3 in their methodology conflated both variables (i.e., English was always the L1, Spanish was always the L2, and the L3 was either French or Italian). This same confound is not true of Bardel and Falk (2007) and Falk and Bardel (2011), so it is interesting that they show a very strong L2 effect despite apparent structural proximities between the L3 and the L1. Nevertheless, a number of studies call into question the absolute position of L2 transfer, thus rendering the steadfast line of the L2 Status Factor problematic (e.g., Na Ranong and Leung, 2009; Hermas, 2010; Iverson, 2010; Rothman, 2010, 2011; Montrul et al., 2011; Giancaspro et al., 2015; Slabakova and García Mayo, 2015).

It might be suggested that L2 transfer even under this approach can be circumvented by structural or other factors, which Bardel and Falk do not deny in their published work (see for example Falk et al., 2015[4]). However, it seems unclear how this would be possible under the current explanation based on a DP difference between the L1 and other grammars and the hypothesized suppression of the L1 that this creates. In other words, it is not clear how or why factors such as relative structural similarity could bypass the filter imposed by purported cognitive differences (reliance on declarative vs. procedural memory) related to the L1 and L2.

---

[4]In a recent paper, Falk et al. (2015) acknowledge that with certain populations typological relatedness might trump the L2 privilege. However, the authors are very clear that such a possibility only obtains in learners that are metalinguistically aware, even trained, in their L1 and L2, for example individuals who are trained teachers of their L1 as well as successful learners of an L2.

The Cumulative Enhancement Model

The CEM proposed by Flynn et al. (2004) posits that both the L1 and the L2 are possible sources of morphosyntactic transfer at the initial stages of L3 acquisition. The CEM maintains that language acquisition is a collective process throughout the lifespan whereby experience with the acquisition of any prior language can facilitate subsequent language acquisition. Differently from the L2 Status Factor, the CEM claims that previous linguistic knowledge transfers in multilingual development from any language available to the learner, irrespective of order of acquisition. However, transfer crucially only obtains when such knowledge has a facilitative effect, since language acquisition is assumed to be a non-redundant process. Alternatively, when transfer from either language would not be facilitative it is effectively blocked.

Flynn et al. (2004) base their claims on data from the production of restrictive relative clauses in L1 Kazakh/L2 Russian/L3 English speakers. Their data demonstrate that experience in any previously acquired language can be taken advantage of, providing support for the CEM. Still, there has not been much published work that supports the CEM unambiguously (but see Jaensch, 2011; Berkes and Flynn, 2012, for claims of support for a 'weak' version of the CEM; see also Slabakova and García Mayo, 2015, for a discussion of the roles of cumulative enhancement and its interaction with cumulative inhibition).

Supported by a growing literature, as we will see in greater detail below, is the CEM's claim that transfer is not restricted to a default L1 or default L2. Amassing evidence in the generative L3 transfer literature supports the CEM's claim that acquisition is inherently non-redundant by cognitive design. Conversely, the strong claim that non-facilitative transfer cannot obtain is simply not supported by much of the available evidence. The evidence reviewed above related to the L2 Status Factor already demonstrates counter evidence to such a claim. Clear motivations for why the CEM rejects non-facilitative transfer as a possibility remain elusive. From our perspective, having to avoid non-facilitative transfer *a priori* places an unrealistic burden on limited cognitive resources during the formation of the L3/Ln system. At a minimum, it implies that the learner would have to have enough experience with the L3/Ln on a property-by-property basis to determine what could be facilitative, and also to suppress what would be non-facilitative even when strong evidence of overall structural similarity between two of the grammars is overwhelming. It also seems to suggest that transfer is incremental throughout L3 development. As such, both the L1 and the L2 would need to remain equally activated throughout the L3 process, which entails a cognitive cost that creates a burden on finite resources.

The Typological Primacy Model

The TPM (Rothman, 2010, 2011, 2013, 2015) is a model of L3/Ln transfer that, similar to the CEM, envisions access to both the L1 and L2 mental grammars at the initial stages. Differently from the CEM, however, the TPM acknowledges the possibility of non-facilitative transfer, which derives from the same general spirit underlying the original CEM: for reasons of general cognitive

economy, language acquisition is forced to be a non-redundant process. Both the CEM and the TPM agree that multilingualism is conditioned by a cumulative effect of previous linguistic acquisition; however, the TPM views selection of a language for transfer as being conditioned by factors related to underlying structural similarity between the languages at play, as opposed to mere facilitation.

Recall that for the CEM, transfer at the initial stages and beyond is predicted to be maximally facilitative or otherwise neutralized. Unlike the CEM, the TPM hypothesizes that transfer is complete (the entire L1 or L2) and early in L3 interlanguage development, and is determined by the structural similarity between the target L3 and the L1 or L2, as assessed by the internal (linguistic) parser. More precisely, it makes reference to structural similarities at an underlying level of linguistic competence across the three languages. Therefore, the possibility of non-facilitative transfer is taken not only to be possible, like the L2 Status Factor (albeit for different reasons), but rather predictable.

Proposals for how the linguistic parser determines at an early stage whether the L1 or L2 should transfer have been the topic of recent work (Rothman, 2013, 2015). Following the logic advocated in Schwartz and Sprouse's (1996) Full Transfer/Full Access Hypothesis for L2 acquisition, the TPM advances the idea that one of the two systems must be transferred completely in the initial stages. A continuum of cues related to four factors is hypothesized to lead the parser to select between the two available grammars, represented in **Figure 1**.

Not all of these factors are as easily usable by or equally accessible to the parser at the same time, partially depending on the specific language pairings. For this reason, the above list is intended to be implicationally hierarchical. The TPM does not idealize an unrealistic situation in which these four factors are mutually exclusive to one another. Rather, there is clear mutual dependency of the levels in the hierarchy. For example, syntactic structure clearly depends on functional morphology, which in

turn is determined in the lexicon and interfaces with phonology. Rothman (2013) makes it clear that, of the four possible types of cues, it is ultimately the language combinations themselves that determine how many and which, if any, of the four factors are usable. Ultimately the TPM predicts that the previously acquired linguistic system with the most detectable/usable structural crossover, at the highest levels of the cue hierarchy, at the earliest of timing at the very initial stages of L3 will be selected for complete transfer.

Now let us turn our attention to the empirical evidence in support of the TPM. Rothman (2010) examined the L3 acquisition of Brazilian Portuguese, contrasting two sets of L3 learners: (a) L1 speakers of English who were highly proficient learners of L2 Spanish and (b) L1 speakers of Spanish who were highly proficient learners of L2 English. The study examined word order restrictions relating to transitive verbs and two types of intransitive verbs (unergatives and unaccusatives) in declaratives and interrogatives, as well as relative clause attachment preference. Despite the fact that Spanish and Brazilian Portuguese are typologically similar, Brazilian Portuguese patterns much more like English than Spanish in these related domains. The data unambiguously show Spanish transfer irrespective of whether it was an L1 or L2, supporting the TPM and providing evidence against the predictions of the L2 Status Factor and the CEM.

In recent years, several studies have shown that relative structural similarity between the L3 and one of the previously acquired systems is the most deterministic factor for multilingual transfer. Much of the additional work supporting the typological factor in adult multilingualism comes from language triads where two Romance languages and English are involved (e.g., Foote, 2009; Iverson, 2009, 2010; Ionin et al., 2011; Montrul et al., 2011; Borg, 2013; Giancaspro et al., 2015). This fact might leave one questioning whether the TPM makes predictions beyond such obvious language pairings in the Romance family (see Rothman, 2015). If the TPM is on the right track, predictions should be derivable irrespective of the languages implicated in any triad. Rothman's (2013, 2015) articulation of the TPM claims that it makes universal predictions. Promisingly, recent research with more varied L3 language pairings has shown similar support for the TPM (e.g., L1 Tuvan/L2 Russian/L3 English, Kulundary and Gabriele, 2012; L1 Uzbek/L2 Russian/L3 Turkish, Özçelik, 2013; L1 Polish/L2 French/L3 English, Wrembel, 2012; L1 English/L2 Spanish/L3 Arabic, Goodenkauf and Herschensohn, 2014).

For example, Özçelik (2013) examined the L3 acquisition of Turkish by Uzbek-Russian bilinguals with respect to quantificational scope. For ease of exposition, we will use English to explain the linguistic facts. Whereas Uzbek (similar to English) has both surface and inverse scope interpretations of sentences like (1), Turkish only has the surface scope interpretation (2).

(1) Jack didn't find two guys.
  ✓ (a) Surface: It is not the case that Jack found two guys. (Jack found one guy, three guys, no guys, etc.)
  ✓ (b) Inverse: There are two guys that Jack didn't find.

FIGURE 1 | Implicational hierarchy of input cues. Adapted from Rothman (2013, 2015).
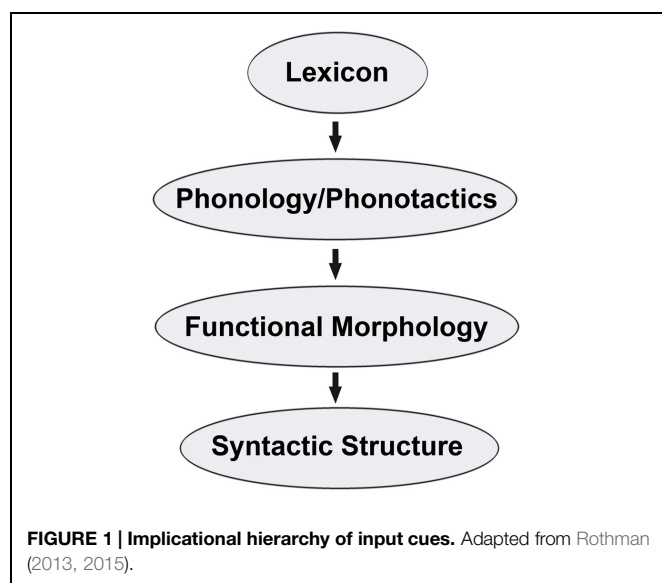
(2) Jack iki kişi bul-ma-dı.
    Jack two person find-NEG-PAST
    "Jack didn't find two guys."
✓ (a) Surface: It is not the case that Jack found two guys. (Jack found one guy, three guys, no guys, etc.)
✗ (b) Inverse: There are two guys that Jack didn't find.

The L3 acquisition of Turkish by Uzbek–Russian bilinguals in this regard is interesting in that, although Turkish and Uzbek are both Turkic languages and are typologically related, Turkish behaves like Russian with respect to this structure, and differently from Uzbek, which allows both scope interpretations. The results show that the learners treat Turkish like Uzbek, as they allow both surface and inverse scope interpretations of sentences like (2), i.e., they transfer from the holistically TYPOLOGICALLY similar language (Uzbek), rather than from Russian, the language that is STRUCTURALLY similar to Turkish for this particular property. Results support the TPM, as transfer is activated on the basis of overall typological similarity, even though this leads to a less optimal grammar since the source language for transfer (Uzbek) and the target language (Turkish) behave differently with respect to the construction tested here and despite the fact that Russian, the L2, would have been more facilitative for this property.

EEG and the ERP Methodology: Use and Application to L3

To date, all of the experimentation done under the current models of L3/Ln transfer has been methodologically behavioral. Although illuminating, we will argue that these models also make predictions that can be tested with online methodologies, such as ERP. We argue that testing these predictions can add new insights to and strengthen the descriptive and explanatory power of these models.

EEG and ERPs

EEG is an electrophysiological method that records at the scalp the electrical activity generated by large populations of neurons firing in synchrony. It provides high temporal resolution, with millisecond precision, and therefore it is an excellent tool to examine the dynamics of language processing as it unfolds over time. However, unlike methods such as functional magnetic resonance imaging (fMRI) or positron emission tomography (PET), EEG provides limited spatial resolution, due to the fact that the signal recorded at the scalp cannot be unambiguously traced back to its source (Friederici, 2004). Event-related potentials (ERPs) are small voltage changes that are time-locked to a specific event of interest. For example, if the event of interest is agreement resolution, we can time-lock the EEG signal to the element in the sentence where the parser can determine whether or not agreement was successful (e.g., *Harold saw this house/*houses yesterday*). If a comparison across conditions (e.g., grammatical vs. ungrammatical) reveals differences in the morphology of the waveforms, we can assume that the brain was sensitive to the property under investigation. One clear advantage of ERPs is their multidimensional nature. ERPs can

be examined in terms of their latency (the time window when the effect emerges), amplitude (the strength of the effect), and polarity (whether the voltage change is negative or positive). They can also be examined in terms of their scalp topography (the electrode region or regions where the effect is captured). Importantly, this allows for a very in-depth characterization of the mechanisms underlying language processing and for a very fine-grained comparison between different populations (e.g., native speakers vs. adult language learners). One of the most unique advantages of the ERP methodology is the fact that different ERP components, such as the N400 and the P600, are modulated by different aspects of language processing. The P600 (e.g., Osterhout and Holcomb, 1992; Hagoort et al., 1993) is a positive deflection between 500 and 900 ms whose elicitation is attributed to processes of syntactic reanalysis (e.g., Osterhout and Holcomb, 1992; Gouvea et al., 2010), syntactic integration (e.g., Kaan et al., 2000), and syntactic repair (Hagoort et al., 1993; Osterhout and Mobley, 1995). While not all processes which affect the P600 are syntactic (or even linguistic) in nature, it is noteworthy that this is the only component that is consistently found for syntactic agreement violations in native speakers (e.g., Coulson et al., 1998; Gunter et al., 2000; Hagoort, 2003; Wicha et al., 2004; Barber and Carreiras, 2005; Martín-Loeches et al., 2006; Nevins et al., 2007; Frenck-Mestre et al., 2008; O'Rourke and Van Petten, 2011), making it the most reliable ERP signature associated with the native processing of syntactic agreement.

In contrast, the N400 is a negative-going wave between 200 and 600 ms which typically emerges in central posterior electrodes of the EEG cap and which has been found to be sensitive to the strength of lexical associations (see Kutas and Federmeier, 2011 for a review). For example, words that are semantically associated with a previously presented prime (e.g., *dog-cat*) show reduced N400 amplitudes relative to words unrelated to the prime (e.g., *car-pen*) (Holcomb and Neville, 1990). Studies on native processing where the only ERP signature associated with syntactic agreement violations is the N400 are rare. One exception is Barber and Carreiras (2005), who examined number and gender violations in Spanish word pairs, and found a larger N400 for both violation types relative to grammatical strings. Since isolated word pairs do not require syntactic structure building, Barber and Carreiras (2005) interpret these findings as evidence that the Spanish native speakers processed the agreement violations at the lexical level, by comparing the lexical features of the agreeing words. Interestingly, when the exact same violations were examined in sentences, they yielded a P600.

In a subset of studies, the P600 is preceded by a negative-going wave in the N400 time window, sometimes with a left anterior distribution. The qualitative nature of this negativity is very much a matter of debate. Some authors have identified it as the Left Anterior Negativity (LAN), a component argued to index automatic morphosyntactic processing (e.g., Friederici et al., 1996). A problem with this interpretation, however, is that a number of studies examining morphosyntactic processing in native speakers do not find the LAN for agreement errors (e.g., Wicha et al., 2004; Frenck-Mestre et al., 2008; Alemán Bañón

et al., 2012). Alternatively, this negativity has been identified as an N400. Under this interpretation, the left anterior distribution of the N400 results from its topographical overlap with a central-posterior P600, which cancels out the negativity in central-posterior regions of the scalp (e.g., Guajardo and Wicha, 2014; Tanner and Van Hell, 2014). Under this view, the N400 is argued to reflect either the semantic integration difficulty caused by the presence of the agreement error (e.g., Guajardo and Wicha, 2014), or individual differences with respect to processing strategies, with some individuals relying on lexical information (N400) and others on combinatorial information (P600) (Tanner, 2013, 2015; Tanner and Van Hell, 2014). Importantly for the purposes of the present study, it is the P600 that consistently emerges for morphosyntactic errors in native speakers, even if sometimes it is preceded by a negativity. The reverse, however, is not true. As stated in Tanner (2015), agreement errors in native speakers are unlikely to yield an N400 not followed by the P600:

> "(. . .) given the dominance of P600 effects in response to morphosyntactic violations across individuals, it is highly unlikely to randomly draw a sample of individuals where only a reliable N400 would be found, with no following P600 — even though some individuals show negativity-dominant brain responses to morphosyntactic violations."
>
> (Tanner, 2015, p.154).

ERP and Formal Linguistic Approaches to SLA

How can we use the ERP methodology to test formal linguistic theoretical models of adult language acquisition? To give one example, Alemán Bañón et al. (2014) relied on the difference between the N400 and the P600 to adjudicate between the Full Transfer/Full Access Hypothesis (Schwartz and Sprouse, 1996) and the Interpretability Hypothesis (Tsimpli and Dimitrakopoulou, 2007; see also Gabriele et al., 2013a). The study examined the processing of number and gender agreement in L2 Spanish by advanced English-speaking learners. Critically, these two hypotheses differ with respect to whether or not adult L2 learners are predicted to be able to show native-like processing for novel uninterpretable features (in this case, Spanish gender agreement). Only the Full Transfer/Full Access Hypothesis predicts so, since L2 acquisition is hypothesized to be influenced but not constrained by the properties of the L1 (e.g., White et al., 2004).

Under the Interpretability Hypothesis, in contrast, English-speaking learners of Spanish are not predicted to show native-like processing for gender agreement, regardless of proficiency. Learners might exhibit behavior that appears native-like (e.g., high accuracy rates in behavioral tasks; see Franceschina, 2005 for an example), but the supporters of the Interpretability Hypothesis argue that such behavior is achieved through compensatory strategies (e.g., Hawkins, 2001). For example, learners might establish associations between morphemes that tend to co-occur, in which case gender violations might yield a larger N400 than grammatical sentences (similar to what Barber and Carreiras, 2005, found for word pairs in Spanish native speakers). Alternatively, learners might rely on the phonological similarity between the agreeing words (in Spanish, most masculine nouns

end in –o and most feminine nouns end in –a), in which case gender violations should only modulate the N400 component, consistent with a number of studies which have examined the effects of phonological similarity on word processing[5].

Alemán Bañón et al.'s (2014) proposal is that if English-speaking learners of Spanish can process novel features in a native-like manner, they should show a P600 for gender violations, consistent with a large body of literature which reports P600 effects for agreement violations in native speakers (including the Spanish-speaking controls reported in Alemán Bañón et al., 2012, 2014, for whom this was the only component found for number and gender violations across the different syntactic contexts tested). However, if learners rely on other mechanisms, such as comparing the lexical features of the agreeing words or relying on their phonological similarity (as would be predicted by the Interpretability Hypothesis), gender violations should yield a larger N400 than grammatical sentences (e.g., Barber and Carreiras, 2005; Coch et al., 2008). The advanced L1 English L2 Spanish learners in Alemán Bañón et al. (2014) showed robust P600 effects (and no N400) for both number and gender violations overall. This evidence was used to argue that native-like processing for features that are unique to the L2 is possible in adult L2 acquisition, consistent with full UG accessibility in adulthood. These results are also consistent with previous ERP studies providing evidence that, at an advanced level of proficiency, adult learners can exhibit native-like processing for L2 morphosyntactic properties (e.g., Rossi et al., 2006), including those that are not instantiated in the L1 (e.g., Dowens et al., 2010, 2011; Foucart and Frenck-Mestre, 2012). What is most relevant about the approach by Alemán Bañón et al. (2014) is that it shows how the ERP methodology can be used to shed light on the qualitative nature of L2 processing and, more importantly for the present discussion, to test current theoretical models of adult language acquisition.

In another relevant study, Bond et al. (2011) found a P600 for both number and gender violations in adult English-speaking learners of Spanish at a lower level of proficiency. Interestingly, the L2 learners also showed a larger P600 for number (present in the L1) than gender (unique to the L2) violations, which is consistent with the possibility that, at lower levels of proficiency, processing is more heavily impacted by L1 transfer (e.g., Tokowicz and MacWhinney, 2005; see Dowens et al., 2010, and Foucart and Frenck-Mestre, 2011, for further evidence for transfer effects in advanced learners).

Importantly for the present discussion, ERP has also been used to examine the initial stages of L2 processing. For example, McLaughlin et al. (2010) tracked L1 English learners throughout their first year of university L2 French. The linguistic focus of the study was subject-verb agreement, which is instantiated in both English and French, and article-noun number agreement, which is only instantiated in French. For subject-verb agreement violations, a subset of "fast" learners ($n = 7$) showed an N400 effect (violations being more negative than grammatical

---

[5]For example, words which are phonologically similar to their prime (e.g., *lake-break*) show a reduction in N400 amplitude compared to words that are phonologically unrelated to the prime (e.g., *lake-line*) (e.g., Coch et al., 2008).

sentences) after only 1 month of instruction, which the authors interpret as evidence that learners were sensitive to the violations but did not process them grammatically from the start. After 4 and 6 months of instruction, however, the same violations yielded a P600 (similar to the native controls). Article-noun number violations, in contrast, did not yield any effects at any point. In light of these results, McLaughlin et al. (2010) argue against full transfer in the initial stages, since learners did not show evidence of grammatical processing for the property that was available through the L1 (subject-verb number). Instead, the authors propose that learners initially treat all grammatical violations at the lexical level by relying on co-occurrence frequencies between morphemes (e.g., pronouns and verbal inflection; see also Ullman, 2001, 2005).

The results by McLaughlin et al. (2010) are not supported by another longitudinal study by Gabriele et al. (2013b). The authors examined morphosyntactic development in novice English-speaking learners of Spanish. The study focused on three types of agreement: (1) subject-verb number, which is realized in both English and Spanish, (2) noun-adjective number, which is only morphologically realized in Spanish, and (3) noun-adjective gender, which is unique to Spanish. In native speakers, all violation types yielded robust P600 effects (Bond et al., 2011). Interestingly, the learners (n = 23) showed a small positivity in the P600 time window for both types of number violations (feature that is present in the L1) after only 2 months of instruction. Crucially, after 6 months of instruction, this positivity became more robust and showed a broader scalp distribution, more in line with the canonical P600 elicited by the Spanish controls. Gender violations, in contrast, yielded neither N400 nor P600 effects at any point. Since the learners showed sensitivity (a positivity) to the feature that is shared by the L1 and L2 (number) after only 2 months of instruction, Gabriele et al. (2013b) argue in support of theories that assign a privileged role to the properties of the L1 at the initial stages.

The above studies provide very relevant findings for our goal of using ERP to examine the initial stages of L3/Ln acquisition. The logic is as follows: if L2ers show ERP signatures akin to native speakers for a given grammatical property, then we can assume that, in principle, the property at stake is available as a source of transfer. If so, we might expect that advanced L1 English L2 Spanish bilinguals learning Portuguese as an L3 might show a positivity in the P600 time window for both number and gender violations in Portuguese. Showing this for gender would make them different from the English-speaking learners of Spanish reported in Gabriele et al. (2013b), who only showed this positivity for number. Such findings would be consistent with the TPM and the CEM (for different reasons), but crucially not with the L2 Status Factor. Recall that, under the current formulation of the L2 Status Factor, the L2 and L3 are hypothesized to be stored in declarative memory. As stated in Ullman (2001, 2005), learners' greater reliance on declarative memory is predicted to yield N400 effects for grammatical violations where native speakers show qualitatively different components (e.g., a biphasic LAN-P600 pattern according to Ullman, 2001). Therefore, if the L2 Status Factor is on the right track, novice learners of L3 Portuguese whose L1 and L2 are English and Spanish,

respectively, should show, at most, N400 effects for gender agreement violations in L3 Portuguese. This is one example of how the ERP methodology (i.e., the fact that the N400 and the P600 have been argued to be associated with different aspects of processing and different memory systems) can be used to adjudicate between the above models in a way that behavioral methodologies cannot. With respect to the CEM and the TPM, since transfer by either facilitation (CEM) or by typological proximity (TPM) would always favor Spanish transfer, there is no way to tease apart these models with the present domain of grammar. In Section "Sample ERP Methodology," we will provide a sample methodology that is able to tease apart all three initial stages models.

Sample ERP Methodology

In order to test the above models of L3 acquisition, we detail a novel methodology that is part of our in progress work, which relies on the use of artificial languages (AL) as L3s and which combines behavioral and processing measures (i.e., grammaticality judgment task and ERP data). The use of ALs offers two crucial advantages. First, we can test truly *ab initio* learners, allowing us to better contrast the predictions of the above models, all of which are initial stages models. Second, by using ALs we can systematically manipulate the similarity between the L3 and the L1/L2 in terms of (1) the presence/absence of a given feature and (2) the levels of the cue hierarchy which, according to Rothman (2013, 2015), will determine the parser's selection of a transfer source. In addition, the use of ERP will shed light on the qualitative nature of processing at L3 initial stages. This is especially relevant, given the current articulations of the L3 models under review. For example, the L2 Status Factor (Bardel and Falk, 2012) argues that L3 acquisition relies mainly on declarative memory and, therefore, L3 beginners are predicted to show N400 responses for morphosyntactic properties associated with qualitatively different components in native speakers (e.g., P600 or a biphasic LAN-P600; e.g., Ullman, 2001; Morgan-Short et al., 2012). In contrast, the TPM assumes that the initial state of L3 acquisition is the entire L1 or L2 and, therefore, this model predicts that "transferable" morphosyntactic properties should be associated with ERP signatures that are qualitatively native-like from the start (e.g., P600; Rothman, 2015).

The linguistic focus of the proposed study is number and gender agreement. This choice is motivated on the basis that most previous ERP studies looking at the initial stages of L2 processing have focused on this domain (e.g., Osterhout et al., 2006; Morgan-Short et al., 2010; Gabriele et al., 2013b). Therefore, we can make predictions regarding the initial stages of L3 processing based on our knowledge of how agreement in processed at the initial stages of L2 acquisition. In addition, our study could provide insight into the differences and similarities between the L2 and L3 acquisition of these grammatical properties. Our rationale is based on two core findings: (1) The longitudinal study by Gabriele et al. (2013b) looking at L1 English beginners of L2 Spanish shows ERP signatures consistent with transfer of grammatical number (present in the learners' L1) from the

earliest of stages tested; (2) A number of studies have shown native-like ERP signatures for grammatical gender in advanced L1 English learners (e.g., Dowens et al., 2010, 2011; Foucart and Frenck-Mestre, 2012; Gabriele et al., 2013a; Alemán Bañón et al., 2014). From (1) we believe it reasonable to use ERP to examine transfer at the initial stages of L3 acquisition. Furthermore, (2) suggests that, for the acquisition of an L3 that realizes gender agreement, we can predict sensitivity to gender not only in L3ers who are L1 Spanish-L2 English, but also in L3ers who are L1 English-L2 Spanish (provided they have reached a high level of proficiency in L2 Spanish). If both groups show sensitivity to grammatical gender in the L3, this would immediately call into question the L2 Status Factor (especially if brain responses are not in the form of N400 effects, which is the component argued to be associated with declarative memory).

Recall, however, that—for the above learning scenario—both the CEM and the TPM predict the transfer of gender irrespective of L1/L2 sequencing. The two models differ in the conditions under which this transfer should happen. Under the TPM, the learner's perceived similarity between the L3 and the L1/L2 will determine the source of transfer. Under the CEM, gender will be transferred when appropriate, based on the fact that it has already been acquired in a previous language (Spanish). Our design contrasts the predictions of these two models by using two ALs as L3s. One of the ALs is lexically similar to English ("Mini-English") and the other one, to Spanish ("Mini-Spanish"), but they both instantiate number and gender agreement. This lexical similarity between English and Mini-English should have a non-facilitative effect under the TPM (i.e., the parser should assume that Mini-English does not instantiate gender based on the fact that English does not realize this property). Under the CEM, this negative transfer should be blocked, and the parser will transfer gender from the facilitative language, Spanish.

Artificial Languages

Following work by Williams and colleagues (e.g., Williams, 2004; Williams and Kuribara, 2008; Marsden et al., 2013), Mini-English is built on the English lexicon and novel morphemes for number and gender have been added to articles and adjectives. The second AL, Mini-Spanish, is based on the Spanish lexicon where also completely novel morphemes for number and gender have been added to articles and adjectives. Each AL includes 12 inanimate nouns (six masculine, six feminine) and 12 adjectives, in order to facilitate the learning of its lexicon. Each AL also includes one article that inflects for number and gender (four variants: masculine-singular, feminine-singular, masculine-plural, feminine-plural), one copulative verb that inflects for number (singular, plural), one conjunction, one adverb, and two locatives. Since one of our research questions concerns the role of lexical similarity on the selection of a transfer source, all other potential cues are neutralized in the ALs. For example, training in the AL will take place in the visual modality (as opposed to aural), to avoid providing phonological information. Likewise, learners will only be exposed to meaningful examples of the AL where word order is similar in English and Spanish, in order to neutralize word order as a cue. Examples of short sentences in Mini-Spanish are provided in (3) and (4) below:

(3)

(a) <u>Ne</u> **camion** es <u>car-enu.</u>
the-$_{MASC\text{-}SG}$ truck is expensive-$_{MASC\text{-}SG}$

(b) <u>Ner</u> **camion** son <u>car-enur.</u>
the-$_{MASC\text{-}PL}$ truck are expensive-$_{MASC\text{-}PL}$

(c) <u>Ge</u> **llave** es <u>car-egu.</u>
the-$_{FEM\text{-}SG}$ key is expensive-$_{FEM\text{-}SG}$

(d) <u>Ger</u> **llave** son <u>car-egur.</u>
the-$_{FEM\text{-}PL}$ key are expensive-$_{FEM\text{-}PL}$

(4)

(a) <u>Ge</u> **llave** es sobre ne reloj.
the key is above the watch.

(b) <u>Ge</u> **llave** es bajo ne reloj.
the key is below the watch

As can be seen in (3a-b), the masculine noun *camion* "truck," which has been selected from the Spanish lexicon, must agree in number and gender with the preceding article (masculine-singular: *ne*; masculine-plural: *ner*) and the predicative adjective (masculine-singular: *carenu*; masculine-plural: *carenur*). A similar example is provided in (3c-d), where the feminine noun *llave* "key," also from the Spanish lexicon, agrees in number and gender with the preceding article (feminine-singular: *ge*; feminine-plural: *ger*) and the predicative adjective (feminine-singular: *caregu*; feminine-plural: *caregur*). All of the nouns in Mini-Spanish have the same lexical gender as their Spanish counterparts. Importantly, all nouns have been selected such that, despite their lexical similarity with their equivalent in Spanish, they do not exhibit the markers typically associated with the masculine/feminine distinction in Spanish (e.g., masculine –o, feminine –a), to avoid providing learners with additional morphological cues. Notice also that, similar to Morgan-Short et al.'s (2010) study, the nouns *camion* and *llave* provide no phonological cues regarding the gender of the noun. This was done in an attempt to prevent learners from relying on a purely phonological strategy when computing gender agreement. In order for the comparison between number and gender to be more ecologically valid, nouns in the ALs are also opaque for number, as shown in (3a-b) and (3c-d). The sentences in (4) show the distribution of the locatives "above" and "below" in Mini-Spanish. With respect to the design of Mini-English, semantically equivalent nouns and adjectives were used (e.g., *truck*, *key*). With respect to lexical gender, since English lacks this property altogether, we decided to assign Mini-English nouns the same lexical gender as the nouns in Mini-Spanish (i.e., *truck* and *key* are masculine and feminine, respectively, similar to *camion* and *llave*). Examples of mini-English are provided in (5) and (6) below:

(5)

(a) <u>Ne</u> **truck** is <u>expens-enu.</u>
the-$_{MASC\text{-}SG}$ truck is expensive-$_{MASC\text{-}SG}$

(b) <u>Ner</u> **truck** are <u>expens-enur.</u>
the-MASC-PL truck are expensive-MASC-PL

(c) <u>Ge</u> **key** is <u>expens-egu.</u>
the-FEM-SG key is expensive-FEM-SG

(d) <u>Ger</u> **key** are <u>expens-egur.</u>
the-FEM-PL key are expensive-FEM-PL

(6)

(a) <u>Ge</u> **key** is above ne watch.
the key is above the watch.

(b) <u>Ge</u> **key** is below ne watch.
the key is below the watch

The structure of interest will be the agreement relation between the noun and the predicative adjective, which will be located across a verb phrase (VP; e.g., *the* <u>truck</u> VP[*is* <u>expensive</u>]). Although it has been argued that agreement relations are more taxing when they are non-local (i.e., across a verb phrase) for both native speakers (e.g., Alemán Bañón et al., 2012) and L2 learners at an advanced level of proficiency (Foucart and Frenck-Mestre, 2012; Alemán Bañón et al., 2014), our choice is motivated upon the grounds that this is a syntactic context where English and Spanish exhibit similar word order (e.g., *el camión es caro* "the truck is expensive"). In contrast, when agreement is local, the position of the adjective with respect to the noun differs in English and Spanish (e.g., *camión caro* "truck expensive"). We are justified in restricting the design of the study to lexical similarity given Rothman's (2013, 2015) claims regarding the primacy of the lexicon for determining transfer [see The Typological Primacy Model (1) above]. Indeed, this is sufficient to test between the three models, which is the primary goal of our study. To further test the very claim of primacy of the lexicon over actual syntactic cues made by the TPM, the next methodological step would be to offer additional competing cues in the ALs. For example, adding to Mini-English a syntactic property that conflicts with the English grammar but is grammatical in Spanish would allow us to test the TPM cue hierarchy independently, since we would have a case where the lexical level is similar to English, but the morphological and syntactic levels are similar to Spanish. The TPM is clear: the lexical level, which is argued to be the most detectable one and, therefore, the top level of the hierarchy, should neutralize the use of the other cues.

Participants

With respect to the participants, our study includes four groups of English-Spanish bilinguals who differ along two criteria: (1) the order of acquisition of English vs. Spanish, and (2) the AL they will be trained on. All L3 learners will have acquired their L2 after ~11 years of age and will have high-proficiency in the L2. After the completion of the L3 study, all learners will be tested in their L2 for knowledge of the relevant properties (i.e., agreement).This is to ensure that the relevant properties are in place in the L2 and can, therefore, transfer to the L3. **Table 1** below offers a schematic of the learner groups in our design.

TABLE 1 | Breakdown of groups based on L1-L2-AL combination.

Group	L1	L2	Languages of training
Group 1 (N = 24)	English	Spanish	Mini-Spanish
Group 2 (N = 24)	English	Spanish	Mini-English
Group 3 (N = 24)	Spanish	English	Mini-Spanish
Group 4 (N = 24)	Spanish	English	Mini-English

Artificial Language Training

The study involves a training session in the AL and a judgment task with an EEG recording. During the training, learners will be exposed to meaningful examples of the AL. No metalinguistic explanations are provided, to ensure training is implicit (e.g., Morgan-Short et al., 2010). The training simulates a picture-sentence matching task (e.g., Mueller et al., 2005). Learners see two pictures showing a contrast (e.g., 3 expensive trucks vs. 3 cheap trucks) and their written description in the AL (e.g., "The trucks are expensive" vs. "The trucks are cheap"). By using both masculine and feminine nouns, both in the singular and in the plural, L3 learners receive implicit input on number and gender agreement between articles, nouns, and adjectives. The training will start with simple article-noun phrases and then move to full sentences like the ones in (3) and (4) above. Filler items will be included which manipulate the location of a noun with respect to another noun, via the locatives "above" and "below." Each noun and adjective is presented an equal number of times throughout the training. The same amount of meaningful examples is provided for number and gender. Learners are exposed to 272 meaningful examples (68 per number/gender combination).

To ensure that learners attend to the training, they will complete a comprehension quiz at the end. Learners see a picture (e.g., 3 cheap trucks) and must select the sentence in the AL that best describes it from among five options. Alongside the correct description of the picture ("The trucks are cheap"), the options include a sentence with a violation of gender agreement, a sentence with a number violation, and a sentence with a double violation (number and gender). In half of the items the violation is realized between the article and the noun and, in the other half, between the noun and the adjective. As a control, the fifth option involves a semantic violation (e.g., "The trucks are expensive"), to ensure that learners are able to extract meaning from the pictures used in the AL training. Filler items involve pictures which manipulate the location of two nouns (e.g., a key above a watch). Here, the possible responses include a sentence that correctly describes the picture ("The key is above the watch"), and four incongruent sentences. Two of the incongruent sentences involve the use of the wrong locative (e.g., "The key is below the watch," "The watch is above the key") and the other two involve the use of incorrect nouns. Upon providing their response, learners receive a "correct" or "incorrect" message, which is visually displayed on the computer screen. No other feedback is provided, to ensure that training in the AL remains as implicit as possible. The quiz includes an equal number of sentences with masculine and feminine nouns, and an equal number of sentences with singular and plural nouns. Each noun and adjective is tested an equal number of times throughout the quiz.

Learners are graduated from the training once they reach above chance accuracy in the quiz, which is defined as the ratio of correct responses to the total number of responses (i.e., 20% accuracy). Learners who score below this threshold must take the training again. This necessarily means that different learners will receive different amounts of training, but it ensures that learners have achieved approximately the same level of proficiency at the time of the EEG recording.

Grammaticality Judgment Task

For the purposes of this task, the 12 nouns in each AL have been crossed with the 12 adjectives, yielding a total of 144 noun-adjective combinations. Those agreement dependencies have been embedded in sentences like the one in (7) below, which has six different versions. The sentence structure where we manipulate agreement is based on a previous study on number and gender agreement in Spanish by Alemán Bañón et al. (2012, 2014). Examples are provided for a sentence with a masculine noun in Mini-Spanish.

(7)

(a) <u>Ne</u> **camion** es <u>car-enu</u> y ne reloj tambien.
the $_{-MASC-SG}$ truck $_{-MASC-SG}$ is expensive $_{-MASC-SG}$ and the watch too

(b) <u>Ne</u> **camion** es *<u>car-enur</u> y ne reloj tambien.
the $_{-MASC-SG}$ truck $_{-MASC-SG}$ is expensive $_{-MASC-PL}$ and the watch too

(c) <u>Ne</u> **camion** es *<u>car-egu</u> y ne reloj tambien.
the $_{-MASC-SG}$ truck $_{-MASC-SG}$ is expensive $_{-FEM-SG}$ and the watch too

(d) <u>Ner</u> **camion** son <u>car-enur</u> y ner reloj tambien
the $_{-MASC-PL}$ truck $_{-MASC-PL}$ are expensive $_{-MASC-PL}$ and the watches too

(e) <u>Ner</u> **camion** son *<u>car-enu</u> y ner reloj tambien
the $_{-MASC-PL}$ truck $_{-MASC-PL}$ are expensive $_{-MASC-SG}$ and the watches too

(f) <u>Ner</u> **camion** son *<u>car-egur</u> y ner reloj tambien
the $_{-MASC-PL}$ truck $_{-MASC-PL}$ are expensive $_{-FEM-PL}$ and the watches too

Each one of the 144 sentences will be assigned to one of three conditions: grammatical (7a,d), number violation (7b,e), or gender violation (7c,f). An equal number of masculine and feminine nouns will be used. Likewise, the study involves an equal number of singular and plural nouns. Learners will read the 144 sentences presented one word at a time using the Rapid Serial Visual Presentation Method (RSVP; SOA: 450/300 ms; Alemán Bañón et al., 2012, 2014) while their brain activity is recorded with EEG. There will be 48 items per condition, which corresponds to the mean number of trials per condition reported in Molinaro et al.'s (2011) review of ERP studies on agreement. As can be seen in (7), the adjective is never sentence-final, to avoid semantic wrap-up effects that have been observed in final position (e.g., Hagoort, 2003). At the end of each trial, learners will perform a grammaticality judgment task (e.g., Mueller et al., 2005; Morgan-Short et al., 2010). The motivation for using a

grammaticality judgment is twofold. First, having information regarding the learners' accuracy will allow us to determine the extent to which learners detected the agreement violations at the behavioral level. Second, it has been argued that the amplitude of the P600 is sensitive to the explicitness of the task. As discussed in Molinaro et al. (2011), the amplitude of the P600 tends to decrease when native speakers are asked to read for meaning, as opposed to focus on grammatical correctness (although it should be noted that the P600 emerges even in the absence of a judgment task; see for example Hagoort et al., 1993). Therefore, since the population of interest involves novice L3 learners, where effects are not predicted to be quantitatively native-like or even robust, we believe it is more appropriate to use a grammaticality judgment task, similar to previous ERP L2 studies using the artificial language paradigm (e.g., Mueller et al., 2005; Morgan-Short et al., 2010).

An additional 96 grammatical fillers will be added to the experimental materials (a total of 240), in order to balance the number of grammatical and ungrammatical sentences in the design. Fillers manipulate the position of a given noun with respect to another noun (see the sentences in 4 and 6 above). Importantly, they do not include adjectives and, therefore, shift the attention away from noun-adjective agreement.

Predictions

All three models predict that all learner groups should show sensitivity to number agreement, since both English and Spanish realize this property. It is for gender agreement that the three models make competing predictions. The L2 Status Factor makes two clear predictions: (1) since only the L2 should transfer, only the learner groups who have Spanish as the L2 (Groups 1 and 2) should show sensitivity to gender violations, even if the L3 being acquired is typologically different from L2 Spanish, as is the case for L1 English-L2 Spanish bilinguals trained in Mini-English; (2) brain responses should index reliance on the declarative memory system across the board, that is, number violations should yield N400 effects (with no evidence of a P600 at this stage) in all groups, and so should gender violations in Groups 1 and 2.

For the CEM, all groups should show qualitatively native-like responses to both number and gender (e.g., P600-like component, similar to the L1 English novice learners of Spanish in Gabriele et al., 2013b, which might be preceded by a negativity) since order of acquisition of Spanish should be inconsequential and such transfer would be facilitative[6]. For the TPM, only the groups who are trained in Mini-Spanish (Groups 1 and 3) should show sensitivity to gender violations, given considerations of

---

[6]As mentioned in Section "EEG and ERPs,", some studies have reported a biphasic N400-P600 pattern for syntactic agreement errors in native speakers, and argued for individual differences in the processing of agreement, with most individuals showing a P600 and a subset of them showing an N400. We have, thus, incorporated in our predictions the possibility that the P600 might be preceded by an N400, but we note that most of the available evidence for individual differences in agreement processing comes from studies which have examined subject-verb agreement with English auxiliary verbs in designs which include lexical semantic violations, which are known to modulate the N400 (e.g., Tanner and Van Hell, 2014; Tanner et al., 2014). It remains an open question whether the same variability might emerge in designs that examine other features (i.e., gender) and other syntactic contexts (noun-adjective agreement), and which do not manipulate semantic congruency.

TABLE 2 | Predicted ERP responses for number and gender agreement.

L1-L2-AL Combination	L2 Status Factor		CEM		TPM	
	Number	Gender	Number	Gender	Number	Gender
L1 English-L2 Spanish, L3: Mini-Spanish	N400	N400	(N400)-P600	(N400)-P600	(N400)-P600	(N400)-P600
L1 English-L2 Spanish, L3: Mini-English	N400	N400	(N400)-P600	(N400)-P600	(N400)-P600	
L1 Spanish-L2 English, L3: Mini-Spanish	N400		(N400)-P600	(N400)-P600	(N400)-P600	(N400)-P600
L1 Spanish-L2 English, L3: Mini-English	N400		(N400)-P600	(N400)-P600	(N400)-P600	

We do not predict quantitatively native-like ERP components for any of the properties under examination in any of the groups (e.g., Gabriele et al., 2013b). We use the terms N400 and P600 to highlight the qualitative differences between the predicted effects. We use parentheses to indicate the possibility that the N400 preceding the P600 for agreement violations under the CEM and the TPM might not emerge.

the typological proximity of the languages. For Groups 2 and 4, the lexical similarity between Mini-English and English should mislead the parser into assuming Mini-English does not realize gender agreement. **Table 2** summarizes the predictions in terms of ERP signatures for number and gender agreement violations for all three models.

Behaviorally, the three models predict that all learner groups should perform above chance levels (i.e., above 50% accuracy) with number agreement, since both English and Spanish realize this property. With respect to gender agreement, the L2 Status Factor predicts that only Groups 1 and 2 (i.e., those with Spanish as the L2) should show above chance accuracy with the detection of gender violations. In contrast, the TPM predicts that only Groups 1 and 3 (i.e., those trained in Mini-Spanish) should show above chance performance with gender violations. Finally, the CEM predicts similar performance for number and gender across all groups.

This example methodology shows how obtaining ERP evidence for the multilingual transfer debate is possible and how its application to the literature dominated by behavioral methodology could add new insights.

Conclusion

In this article, we hope to have shown how the ERP methodology can be used to further our understanding of the factors which impact multilingual transfer. After introducing the main theoretical models of L3 acquisition, we provided relevant evidence from existing ERP studies on the native and non-native processing of agreement which strongly motivates the use of ERP to examine transfer at the initial stages of L3 acquisition (i.e., the central question in all three models discussed). Most importantly, we articulated a methodology from our in progress work which combines the ERP methodology and the artificial language paradigm to examine L3 initial stages transfer and whose novelty resides in the fact that it can adjudicate between current articulations of the L2 Status Factor, the CEM, and the TPM in a way that behavioral methodologies cannot. Here, we focused on the domain of grammatical agreement, but it should be noted that the methodology can also be used to examine other domains of grammar, including those which have been investigated in previous L3 behavioral studies (e.g., word order). Enlightening as it is, evidence for and against the L2 Status Factor, the CEM and the TPM consists exclusively of offline,

behavioral data. Ideally, data from online methodologies, such as ERP, will complement what has been shown behaviorally and add new insights to these models. Corroborative or contradicting evidence from processing can strengthen the descriptive and explanatory power of these models or present novel data requiring refinements to them.

Author Contributions

Funding

Acknowledgments

References

Alemán Bañón, J., Fiorentino, R., and Gabriele, A. (2012). The processing of number and gender agreement in Spanish: an event-related potential investigation of the effects of structural distance. *Brain Res.* 1456, 49–63. doi: 10.1016/j.brainres.2012.03.057

Alemán Bañón, J., Fiorentino, R., and Gabriele, A. (2014). Morphosyntactic processing in advanced second language (L2) learners: an event-related potential investigation of the effects of L1-L2 similarity and structural distance. *Second Lang. Res.* 30, 275–306. doi: 10.1177/0267658313515671

Barber, H., and Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: an ERP comparison. *J. Cogn. Neurosci.* 17, 137–153. doi: 10.1162/0898929052880101

Bardel, C., and Falk, Y. (2007). The role of the second language in third language acquisition: the case of Germanic syntax. *Second Lang. Res.* 23, 459–484. doi: 10.1177/0267658307080557

Bardel, C., and Falk, Y. (2012). "The L2 status factor and the declarative/procedural distinction," in *Third Language Acquisition in Adulthood*, eds J. Cabrelli Amaro, S. Flynn, and J. Rothman (Amsterdam: John Benjamins), 61–78. doi: 10.1075/sibil.46.06bar

Berkes, É., and Flynn, S. (2012). "Further evidence in support of the cumulative-enhancement model: CP structure development," in *Third Language Acquisition in Adulthood*, eds J. Cabrelli Amaro, S. Flynn, and J. Rothman (Amsterdam: John Benjamins), 143–164. doi: 10.1075/sibil.46.11ber

Bond, K., Gabriele, A., Fiorentino, R., and Alemán Bañón, J. (2011). "Individual differences and the role of the L1 in L2 processing: an ERP investigation," in *Proceedings of the 11th Generative Approaches to Second Language Acquisition Conference*, eds J. Herschensohn and D. Tanner (Somerville, MA: Cascadilla Proceedings Project), 17–30.

Borg, K. (2013). "The acquisition of future of probability in L3 Spanish," in *Proceedings of the 12th Generative Approaches to Second Language Acquisition Conference*, eds T. Judy and D. Pascual y Cabo (Somerville, MA: Cascadilla Proceedings Project), 11–21.

Cabrelli Amaro, J. (2013). *The Phonological Permeability Hypothesis: Measuring Regressive L3 Interference to Test L1, and L2 Phonological Representations*. Ph.D. thesis, University of Florida, Gainesville, FL.

Cabrelli Amaro, J., and Rothman, J. (2010). L3 acquisition and phonological permeability: a new test case for debates on the mental representation of non-native phonological systems. *IRAL Int. Rev. Appl. Linguist. Lang. Teach.* 48, 273–294. doi: 10.1515/iral.2010.012

Coch, D., Hart, T., and Mitra, P. (2008). Three kinds of rhymes: an ERP study. *Brain Lang.* 104, 230–243. doi: 10.1016/j.bandl.2007.06.003

Coulson, S., King, J. W., and Kutas, M. (1998). Expect the unexpected: event-related brain response to morphosyntactic violations. *Lang. Cogn. Process.* 13, 21–58. doi: 10.1080/016909698386582

De Angelis, G. (2007). *Third or Additional Language Acquisition*. Clevedon: Multilingual Matters.

de Bot, K., and Jaensch, C. (2015). What is special about L3 processing? *Biling. Lang. Cogn.* 18, 130–144. doi: 10.1017/S1366728913000448

Dowens, G. M., Guo, T., Guo, J., Barber, H., and Carreiras, M. (2011). Gender and number processing in Chinese learners of Spanish: evidence from event related potentials. *Neuropsychologia* 49, 1651–1659. doi: 10.1016/j.neuropsychologia.2011.02.034

Dowens, G. M., Vergara, M., Barber, H., and Carreiras, M. (2010). Morphosyntactic processing in late second-language learners. *J. Cogn. Neurosci.* 22, 1870–1887. doi: 10.1162/jocn.2009.21304

Edwards, M., and Dewaele, J.-M. (2007). Trilingual conversations: a window into multicompetence. *Int. J. Biling.* 11, 221–242. doi: 10.1177/13670069070110020401

Falk, Y., and Bardel, C. (2011). Object pronouns in German L3 syntax: evidence for the L2 status factor. *Second Lang. Res.* 27, 59–82. doi: 10.1177/0267658310386647

Falk, Y., Lindqvist, C., and Bardel, C. (2015). The role of L1 explicit metalinguistic knowledge in L3 oral production at the initial state. *Biling. (Camb. Engl.)* 18, 227–235. doi: 10.1017/S1366728913000552

Flynn, S., Foley, C., and Vinnitskaya, I. (2004). The Cumulative-Enhancement model for language acquisition: comparing adults' and children's patterns of development in first, second and third language acquisition. *Int. J. Multiling.* 1, 3–17. doi: 10.1080/14790710408668175

Foote, R. (2009). "Transfer and L3 acquisition: the role of typology," in *Third Language Acquisition and Universal Grammar*, ed. Y. I. Leung (Clevedon: Multilingual Matters), 89–114.

Foucart, A., and Frenck-Mestre, C. (2011). Grammatical gender processing in L2: electrophysiological evidence of the effect of L1-L2 syntactic similarity. *Biling. Lang. Cogn.* 14, 379–399. doi: 10.1017/S136672891000012X

Foucart, A., and Frenck-Mestre, C. (2012). Can late L2 learners acquire new grammatical features? evidence from ERPs and eye-tracking. *J. Mem. Lang.* 66, 226–248. doi: 10.1016/j.jml.2011.07.007

Franceschina, F. (2005). *Fossilized Second Language Grammars: The Acquisition of Grammatical Gender*. Amsterdam: John Benjamins.

Frenck-Mestre, C., Osterhout, L., McLaughlin, J., and Foucart, A. (2008). The effect of phonological realization of inflectional morphology on verbal agreement in French: evidence from ERPs. *Acta Psychol.* 128, 528–536. doi: 10.1016/j.actpsy.2007.12.007

Friederici, A. D. (2004). Event-related brain potential studies in language. *Curr. Neurol. Neurosci. Rep.* 4, 466–470. doi: 10.1007/s11910-004-0070-0

Friederici, A. D., Hahne, A., and Mecklinger, A. (1996). Temporal structure of syntactic parsing: early and late event-related brain potential effects. *J. Exp. Psychol. Learn. Mem. Cogn.* 22, 1219–1248. doi: 10.1037/0278-7393.22.5.1219

Gabriele, A., Fiorentino, R., and Alemán Bañón, J. (2013a). Examining second language development using event-related potentials: a cross-sectional study on the processing of gender and number agreement. *Linguist. Approaches Biling.* 3, 213–232. doi: 10.1075/lab.3.2.04gab

Gabriele, A., Fiorentino, R., Alemán Bañón, J., and Rossomondo, A. (2013b). Tracking brain responses to morphosyntax: a longitudinal study of novice learners of Spanish. *Paper Presented at the 38th Boston University Conference on Language Development* (*BUCLD 38*). Boston, MA.

García Mayo, M. P., Lázaro Ibarrola, A., and Liceras, J. (2005). Placeholders in the English interlanguage of bilingual (Basque/Spanish) children. *Lang. Learn.* 55, 445–489. doi: 10.1111/j.0023-8333.2005.00312.x

García Mayo, M. P., and Rothman, J. (2012). "L3 morphosyntax in the generative tradition: from the initial state and beyond," in *Third Language Acquisition in Adulthood*, eds J. Cabrelli Amaro, S. Flynn, and J. Rothman (Amsterdam: John Benjamins), 9–32. doi: 10.1075/sibil.46.04pil

García Mayo, M. P., and Villarreal Olaizola, I. (2011). The development of suppletive and affixal tense and agreement morphemes in the L3 English of Basque-Spanish bilinguals. *Second Lang. Res.* 27, 129–149. doi: 10.1177/0267658310386523

Giancaspro, D., Halloran, B., and Iverson, M. (2015). Examining L3 transfer: the acquisition of differential object marking in L3 Brazilian Portuguese. *Biling. Lang. Cogn.* 18, 191–207. doi: 10.1017/S1366728914000339

Goodenkauf, J., and Herschensohn, J. (2014). Gender feature transfer from L2 Spanish to L3 Arabic. *Paper Presented at the European Second Language Association* 24, University of York, UK.

Gouvea, A. C., Phillips, C., Kazanina, N., and Poeppel, D. (2010). The linguistic processes underlying the P600. *Lang. Cogn. Process.* 25, 140–188. doi: 10.1080/01690960902965951

Guajardo, L. F., and Wicha, N. Y. (2014). Morphosyntax can modulate the N400 component: event related potentials to gender-marked post-nominal adjectives. *Neuroimage* 91, 262–272. doi: 10.1016/j.neuroimage.2013.09.077

Gunter, T. C., Friederici, A. D., and Schriefers, H. (2000). Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *J. Cogn. Neurosci.* 12, 556–568. doi: 10.1162/089892900562336

Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *J. Cogn. Neurosci.* 15, 883–899. doi: 10.1162/089892903322370807

Hagoort, P., Brown, C. M., and Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Lang. Cogn. Process.* 8, 439–484. doi: 10.1080/01690969308407585

Hawkins, R. (2001). The theoretical significance of Universal Grammar in SLA. *Second Lang. Res.* 17, 345–367. doi: 10.1191/026765801681495868

Hermas, A. (2010). Language acquisition as computational resetting: verb movement in L3 initial state. *Int. J. Multiling.* 7, 343–362. doi: 10.1080/14790718.2010.487941

Holcomb, P. J., and Neville, H. J. (1990). Auditory and visual semantic priming in lexical decision: a comparison using event-related brain potentials. *Lang. Cogn. Process.* 5, 281–312. doi: 10.1080/01690969008407065

Ionin, T., Montrul, S., and Santos, H. (2011). An experimental investigation of the expression of genericity in English, Spanish and Brazilian Portuguese. *Lingua* 121, 963–985. doi: 10.1016/j.lingua.2010.12.008

Iverson, M. (2009). "N-drop at the initial state of L3 Portuguese: comparing simultaneous and additive bilinguals of English/Spanish," in *Minimalist Inquiries into Child and Adult Language Acquisition: Case Studies Across Portuguese*, eds A. Pires and J. Rothman (Berlin and New York, NY: Mouton DeGruyter), 221–244. doi: 10.1515/9783110215359.2.221

Iverson, M. (2010). Informing the age of acquisition debate: l3 as a litmus test. *Int. Rev. Appl. Linguist.* 48, 219–241. doi: 10.1515/iral.2010.010

Jaensch, C. (2011). L3 acquisition of German adjectival inflection. *Second Lang. Res.* 27, 83–105. doi: 10.1177/0267658310386646

Kaan, E., Harris, A., Gibson, E., and Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Lang. Cogn. Process.* 15, 159–201. doi: 10.1080/016909600386084

Klein, E. C. (1995). "Second versus third language acquisition: is there a difference?" *Lang. Learn.* 45, 419–465. doi: 10.1111/j.1467-1770.1995.tb00448.x

Kulundary, V., and Gabriele, A. (2012). "Examining the role of L2 syntactic development in L3 acquisition," in *Third Language Acquisition in Adulthood*, eds J. Cabrelli Amaro, S. Flynn, and J. Rothman (Amsterdam: John Benjamins), 195–222. doi: 10.1075/sibil.46.13kul

Kutas, M., and Federmeier, K. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annu. Rev. Psychol.* 62, 621–647. doi: 10.1146/annurev.psych.093008.131123

Leung, Y. I. (2007). Second language (L2) and third language (L3) French article acquisition by native speakers of Cantonese. *Int. J. Multiling.* 4, 117–149. doi: 10.2167/ijm041.0

Marsden, E., Williams, J. N., and Liu, X. (2013). Learning novel morphology: the role of meaning and orientation of attention at initial exposure. *Stud. Second Lang. Acquis.* 35, 1–36. doi: 10.1017/S0272263113000296

Martín-Loeches, M., Nigbur, R., Casado, P., Hohlfeld, A., and Sommer, W. (2006). Semantics prevalence over syntax during sentence processing: a brain potential study of noun-adjective agreement in Spanish. *Brain Res.* 1093, 178–189. doi: 10.1016/j.brainres.2006.03.094

McLaughlin, J., Tanner, D., Pitkänen, I., Frenck-Mestre, C., Inoue, K., Valentine, G., et al. (2010). Brain potentials reveal discrete stages of L2 grammatical learning. *Lang. Learn.* 60(Suppl. 2), 123–150. doi: 10.1111/j.1467-9922.2010.00604.x

Molinaro, N., Barber, H., and Carreiras, M. (2011). Grammatical agreement processing in reading: ERP findings and future directions. *Cortex* 97, 908–930. doi: 10.1016/j.cortex.2011.02.019

Montrul, S., Dias, R., and Santos, H. (2011). Clitics and object expression in the L3 acquisition of Brazilian Portuguese: structural similarity matters for transfer. *Second Lang. Res.* 27, 21–58. doi: 10.1177/0267658310386649

Morgan-Short, K., Finger, I., Grey, S., and Ullman, M. T. (2012). Second language processing shows increased native-like neural responses after months of no exposure. *PLoS ONE* 7: e32974. doi: 10.1371/journal.pone.0032974

Morgan-Short, K., Sanz, C., Steinhauer, K., and Ullman, M. (2010). Second language learning of gender agreement in explicit and implicit training conditions: an event-related potential study. *Lang. Learn.* 60, 154–193. doi: 10.1111/j.1467-9922.2009.00554.x

Mueller, J. L., Hahne, A., Fujii, Y., and Friederici, A. D. (2005). Native and nonnative speakers' processing of a miniature version of Japanese as revealed by ERPs. *J. Cogn. Neurosci.* 17, 1229–1244. doi: 10.1162/0898929055002463

Na Ranong, S., and Leung, Y. I. (2009). "Null objects in L1 Thai-L2 English-L3 Chinese: an empirical take on a theoretical problem," in *Third Language Acquisition and Universal Grammar*, ed. Y. I. Leung (Bristol: Multilingual Matters), 162–191.

Nevins, A., Dillon, B., Malhotra, S., and Phillips, C. (2007). The role of feature-number and feature-type in processing Hindi verb agreement violations. *Brain Res.* 1164, 81–94. doi: 10.1016/j.brainres.2007.05.058

O'Rourke, P. L., and Van Petten, C. (2011). Morphological agreement at a distance: dissociation between early and late components of the event-related brain potential. *Brain Res.* 1392, 62–79. doi: 10.1016/j.brainres.2011.03.071

Osterhout, L., and Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *J. Mem. Lang.* 31, 785–806. doi: 10.1016/0749-596X(92)90039-Z

Osterhout, L., McLaughlin, J., Pitkänen, I., Frenck-Mestre, C., and Molinaro, N. (2006). Novice learners, longitudinal designs, and event-related potentials: a means for exploring the neurocognition of second language processing. *Lang. Learn.* 56, 199–230. doi: 10.1111/j.1467-9922.2006.00361.x

Osterhout, L., and Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *J. Mem. Lang.* 34, 739–773. doi: 10.1006/jmla.1995.1033

Özçelik, Ö. (2013). Selectivity in L3 transfer: effects of typological and linguistic similarity in the L3 Turkish of Uzbek-Russian bilinguals. *Paper Presented at the 36th Conference of Generative Linguistics in the Old World* (GLOW 36), Lund.

Paradis, M. (2004). *A Neurolinguistic Theory of Bilingualism* [*Studies in Bilingualism, 18*]. Amsterdam: John Benjamins.

Paradis, M. (2009). *Declarative and Procedural Determinants of Second Languages* [*Studies in Bilingualism, 40*]. Amsterdam: John Benjamins. doi: 10.1075/sibil.40

Rossi, S., Gugler, M. F., Friederici, A. D., and Hahne, A. (2006). The impact of proficiency on syntactic second-language processing of German and Italian: evidence from event-related potentials. *J. Cogn. Neurosci.* 18, 2030–2048. doi: 10.1162/jocn.2006.18.12.2030

Rothman, J. (2010). On the typological economy of syntactic transfer: word order and relative clause attachment preference in L3 Brazilian Portuguese. *Int. Rev. Appl. Linguist.* 48, 245–273. doi: 10.1515/iral.2010.011

Rothman, J. (2011). L3 syntactic transfer selectivity and typological determinacy: the typological primacy model. *Second Lang. Res.* 27, 107–127. doi: 10.1177/0267658310386439

Rothman, J. (2013). "Cognitive economy, non-redundancy and typological primacy in L3 acquisition: evidence from initial stages of L3 Romance," in *Romance Languages and Linguistic Theory 2011*, eds S. Baauw, F. Drijkoningen, L. Meroni, and M. Pinto (Amsterdam: John Benjamins), 217–248. doi: 10.1075/rllt.5.11rot

Rothman, J. (2015). Linguistic and cognitive motivations for the typological primacy model of third language (L3) transfer: timing of acquisition and proficiency considered. *Biling. Lang. Cogn.* 18, 179–190. doi: 10.1017/S136672891300059X

Rothman, J., and Cabrelli Amaro, J. (2010). What variables condition syntactic transfer? A look at the L3 initial state. *Second Lang. Res.* 26, 189–218. doi: 10.1177/0267658309349410

Rothman, J., Cabrelli Amaro, J., and de Bot, K. (2013). "Third language (L3) acquisition," in *The Cambridge Handbook of Second Language Acquisition*, eds J. Herschensohn and M. Young-Scholten (Cambridge: Cambridge University Press), 372–393. doi: 10.1017/CBO9781139051729.023

Rothman, J., Iverson, M., and Judy, T. (2011). Some notes on the generative study of L3 acquisition. *Second Lang. Res.* 27, 5–19. doi: 10.1177/0267658310386443

Schwartz, B., and Sprouse, R. (1996). L2 cognitive states and the full transfer/full access model. *Second Lang. Res.* 12, 40–72. doi: 10.1177/026765839601200103

Slabakova, R., and García Mayo, M. P. (2015). The L3 syntax-discourse interface. *Biling. Lang. Cogn.* 18, 208–226. doi: 10.1017/S1366728913000369

Tanner, D. (2013). Individual differences and streams of processing. *Linguistic Approaches to Bilingualism* 3, 350–356. doi: 10.1075/lab.3.3.14tan

Tanner, D. (2015). On the left anterior negativity (LAN) in electrophysiological studies of syntactic agreement: a commentary on "Grammatical agreement processing in reading: ERP findings and future directions" by Molinaro et al., 2014. *Cortex* 66, 149–155. doi: 10.1016/j.cortex.2014.04.007

Tanner, D., Inoue, K., and Osterhout, L. (2014). Brain-based individual differences in online L2 grammatical comprehension. *Biling. Lang. Cogn.* 17, 277–293. doi: 10.1017/S1366728913000370

Tanner, D., and Van Hell, J. (2014). ERPs reveal individual differences in syntactic processing. *Neuropsychologia* 56, 289–301. doi: 10.1016/j.neuropsychologia.2014.02.002

Tokowicz, N., and MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: an event-related potential investigation. *Stud. Second Lang. Acquis.* 27, 173–204. doi: 10.1017/S0272263105050102

Tsimpli, I., and Dimitrakopoulou, M. (2007). The interpretability hypothesis: evidence from wh-interrogatives in second language acquisition. *Second Lang. Res.* 23, 215–242. doi: 10.1177/0267658307076546

Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: the Declarative/Procedural model. *Biling. Lang. Cogn.* 4, 105–22. doi: 10.1017/S1366728901000220

Ullman, M. T. (2005). "A cognitive neuroscience perspective on second language acquisition: the Declarative/Procedural Model," in *Mind and Context in Adult Second Language Acquisition: Methods, Theory, and Practice*, ed. C. Sanz (Washington, DC: Georgetown University Press), 141–178.

White, L., Valenzuela, E., Koslowska-Macgregor, M., and Leung, Y. I. (2004). Gender and number agreement in nonnative Spanish. *Appl. Psycholinguist.* 25, 105–133. doi: 10.1017/S0142716404001067

Wicha, N. Y., Moreno, E. M., and Kutas, M. (2004). Anticipating words and their gender: an event-related brain potential study of semantic integration, gender expectancy and gender agreement in Spanish sentence reading. *J. Cogn. Neurosci.* 16, 1272–1288. doi: 10.1162/0898929041920487

Williams, J. N. (2004). "Implicit learning of form-meaning connections," in *Form-Meaning Connections in SLA*, eds B. Van Patten, J. Williams, S. Rott, and M. Overstreet (Mahwah, NJ: Erlbaum), 203–218.

Williams, J. N., and Kuribara, C. (2008). Comparing a nativist and emergentist approach to the initial stage of SLA: an investigation of Japanese scrambling. *Lingua* 118, 522–553. doi: 10.1016/j.lingua.2007.03.003

Wrembel, M. (2012). "Foreign accentedness in third language Acquisition; the case of L3 English," in *Third Language Acquisition in Adulthood*, eds J. Cabrelli Amaro, S. Flynn, and J. Rothman (Amsterdam: John Benjamins), 281–309. doi: 10.1075/sibil.46.16wre

The Gradience of Multilingualism in Typical and Impaired Language Development: Positioning Bilectalism within Comparative Bilingualism

Kleanthes K. Grohmann[1,2]* and Maria Kambanaros[2,3]

[1] Department of English Studies, University of Cyprus, Nicosia, Cyprus, [2] Cyprus Acquisition Team, Nicosia, Cyprus, [3] Department of Rehabilitation Sciences, Cyprus University of Technology, Limassol, Cyprus

A multitude of factors characterizes bi- and multilingual compared to monolingual language acquisition. Two of the most prominent viewpoints have recently been put in perspective and enriched by a third (Tsimpli, 2014): age of onset of children's exposure to their native languages, the role of the input they receive, and the timing in monolingual first language development of the phenomena examined in bi- and multilingual children's performance. This article picks up a fourth potential factor (Grohmann, 2014b): language proximity, that is, the closeness between the two or more grammars a multilingual child acquires. It is a first attempt to flesh out the proposed gradient scale of multilingualism within the approach dubbed "comparative bilingualism." The empirical part of this project comes from three types of research: (i) the acquisition and subsequent development of pronominal object clitic placement in two closely related varieties of Greek by bilectal, binational, bilingual, and multilingual children; (ii) the performance on executive control tasks by monolingual, bilectal, and bi- or multilingual children; and (iii) the role of comparative bilingualism in children with a developmental language impairment for both the diagnosis and subsequent treatment as well as the possible avoidance or weakening of how language impairment presents.

Keywords: biolinguistics, clitics, comparative linguality, dialect, executive control, Greek, specific language impairment, socio-syntax

INTRODUCTION

Language acquisition in the multicultural, multiethnic, and especially multilingual environments in which children grow up more and more frequently needs to be paid, correspondingly, closer attention to. This much needed attention concerns a range of educational and sociological issues, just as it is relevant for all matters related to language assessment: determining milestones in typically developing children's language development, assessing problems with language growth early on, diagnosing language impairment, and subsequently developing appropriate speech–language therapy and intervention. Beyond these practical needs, there is also a larger research interest in multilingual acquisition that allows a better view into the underlying cognitive structures.

From the earliest studies of language development, it has become very clear that monolingual language acquisition differs greatly from bi- and multilingual language acquisition—despite fundamental similarities. Depending on where one sets the boundaries, it might even be held that

monolingualism does not really exist, *sensu stricto* (think of different sociolects, idiolects, and so on that every speaker commands). This said, the multilingual child faces a number of obstacles that do not factor into monolingual mother tongue acquisition. Two obvious and well studied factors are the age of onset of children's exposure to each of their two or more native languages and the role, in terms of quantity and quality, of the input they receive in each (e.g., Meisel, 2009; Genesee et al., 2011; Unsworth et al., 2014). In addition, Tsimpli (2014) suggests that the timing in monolingual first language development of the phenomena examined in bi- and multilingual children's performance influences whether a particular linguistic phenomenon is acquired (very) early or late. One aspect explored in the present article is a potential fourth factor (Grohmann, 2014b): language proximity, that is, the closeness between the two or more grammars a multilingual child acquires.

Since this article reports research carried out in Cyprus with local acquirers, we will set the scene by briefly laying out the notion of language proximity as relevant for the context of Greek-speaking Cyprus. In the following, we aim to flesh out the proposed gradient scale of multilingualism within the approach dubbed "comparative bilingualism." The empirical part of this research comes from three types of research. We first report data collected on the acquisition and subsequent development of object clitic placement in the two varieties of Greek spoken in Cyprus by bilectal, binational, bilingual, and multilingual children. The second study draws from the performance on executive control tasks by monolingual, bilectal, and bi- or multilingual children. And finally we address a third line of inquiry on the issue of comparative bilingualism, vis-à-vis multilingual language acquisition: the role of bilingualism in children with developmental language impairment, where we will also briefly consider the diagnosis and subsequent treatment of multilingual (language-)developmentally impaired children. Couched within a biolinguistic outlook to language growth, the research agenda sketched here will eventually offer the opportunity to study the neurobiology of language in different (multi)lingual individuals at different ages. This will be reflected in the Discussion and Outlook, which returns full circle to the idea of "comparative bilingualism" by first extending it further (qua a gradient scale of *multi*lingualism), then connecting it to cognitive neuroscientifically relevant research within the new research area of "comparative biolinguistics" (phenotypic variation such as different manifestations of language impairment and breakdown), and finally suggesting a more holistic agenda for future research investigations: "comparative linguality."

APPROACHING LANGUAGE PROXIMITY FOR LANGUAGE ACQUISITION IN CYPRUS

We begin by echoing Grohmann's (2014b) suggestion of a fourth factor for multilingual language development: language proximity. In fact, the present article builds on Grohmann (2014b), a brief commentary on the epistemological paper by Tsimpli (2014), filling in some details and expanding on others.

With respect to proximity, considering the linguistic closeness or distance between the grammars of all languages a bi- or multilingual child acquires will then allow further entertaining the notion of *comparative bilingualism*. The larger research agenda is one in which comparable phenomena are systematically investigated across bi- and multilingual populations with different language combinations, ideally arranged according to purely structural/grammatical, language typological, or perhaps even areal proximity (e.g., a large body of research in the wake of Thomason and Kaufman, 1988). This is a much larger research project for which "language proximity" first has to be properly defined, which we will leave for future considerations; we are grateful to the reviewers for fruitful discussion and constructive feedback on this issue. It will also have to be decided whether the same measurements of proximity are relevant for bi-/multilingual first language acquisition (Barac and Bialystok, 2012) as it has been argued to apply for second language acquisition (Bialystok, 1997; Birdsong and Molis, 2001) and learning (Ringbom, 2006; Ringbom and Jarvis, 2009; Ceñoz and Gorter, 2011), third language acquisition (see Falk and Bardell, 2010, for an overview), especially beyond much studied phonological influence (Llama et al., 2009; Marx and Mehlhorn, 2010), attrition (Montrul, 2008), or, further removed from acquisition factors, for other aspects of language contact (Thomason, 2001; Aikhenvald, 2007; Jarvis and Pavlenko, 2008).

Our present contribution pursues a much more graspable goal, however, namely to compare different populations of Greek speakers on the same linguistic and non-linguistic tools. These include lexical and morphosyntactic tasks, but also measures on language proficiency, pragmatics, and especially executive control. Our populations range from monolingual children growing up in Greece to multilingual children growing up in Cyprus, with several "shades" in between, all centered around the closeness between the language of Greece (Demotic Greek, typically referred to by linguists as Standard Modern Greek) and the native variety of Greek spoken in Cyprus (Cypriot Greek, which itself comes in different flavors ranging from basi- to acrolect). Detailed family and language history background information was also collected for all participants.

The official language of Greek-speaking Cyprus is Standard Modern Greek (SMG), while the everyday language, hence the variety acquired natively by Greek Cypriots, is Cypriot Greek (CG). Calling CG a dialect of SMG as opposed to treating it as a different language is largely a political question; the proximity between the two is very high, and obviously so: The two varieties largely share a common lexicon, sound structure, morphological rule system, and syntactic grammar. According to *Ethnologue* (Lewis et al., 2015), the lexical similarity between CG and SMG lies in the range of 84–93%, which the authors present as follows (http://www.ethnologue.com/ethno_docs/ introduction.asp): "Lexical similarity can be used to evaluate the degree of genetic relationship between two languages. Percentages higher than 85% usually indicate that the two languages being compared are likely to be related dialects." It yet remains to be seen, however, what the exact criteria for such "lexical similarity" are, and whether the conclusions drawn also extend to grammatical aspects of the linguistic varieties

compared. Of immediate relevance is simply the possibility which the lower bound of this purported similarity allows, namely that, at just below the "clearer" cut-off point of 85%, it is not unambiguously evident that CG should exclusively be treated as a dialect of SMG. (We concur with a reviewer who pointed out that such measurements only indicate that CG and SMG may be dialects/varieties of Modern Greek; this much is surely undisputed).

CG and SMG also differ in each of these levels of linguistic analysis as well—and at times quite substantially so (for a recent in-depth discussion, e.g., Tsiplakou, 2014). To briefly illustrate, there are naturally numerous lexical differences, as expected in any pair of closely related varieties, such as the CG feminine-marked *korua* instead of SMG neuter *koritzi* "girl." Phonetically, CG possesses palato-alveolar consonants, in contrast to SMG, so SMG [cεʹɾɔs] becomes CG [tʃεʹɾɔs] for *keros* "weather." The two varieties use a different morpheme to mark 3rd person plural in present and past tenses, such as CG *pezusin* and *epezasin* instead of SMG *pezun* "they play" and *epezan* "they were playing." On the syntactic level, SMG expresses focus by fronting to the clausal left periphery, while CG employs a cleft-like structure, which it also extensively uses in the formation of *wh*-questions. And there are even pragmatic differences such as in politeness strategies: The extensive use of diminutives in SMG is considered exaggerated by CG speakers. See, among many others, Muller (2002), Grohmann et al. (2006), Terkourafi (2007), Grohmann (2009), Arvaniti (2010), and Tsiplakou (2014) for recent discussions and further references.

Traditionally, Greek-speaking Cyprus is considered a language situation of diglossia between the sociolinguistic L(ow)-variety CG and the H(igh)-variety SMG (Newton, 1972 and much work since, building on Ferguson, 1959 [2003]; for recent overviews, see Arvaniti, 2010; Hadjioannou et al., 2011; Rowe and Grohmann, 2013). Moreover, while there is a clear basilect ("village Cypriot"), there are arguably further mesolects ranging all the way up to a widely assumed acrolect ("urban Cypriot"); Arvaniti (2010) labeled the latter Cypriot Standard Greek (CSG), a high version of CG which is closest to SMG among all CG lects. In fact, such CSG may be the real H-variety on the island, on the assumption that without native acquirers of SMG proper, the only Demotic Greek-like variety that could be taught in schools is a "Cyprified Greek," possibly this ostensible yet elusive CSG. However, SMG can be widely heard and read in all kinds of media outlets, especially those coming from the Hellenic Republic of Greece. Note also that there is still no grammar of CSG available, no compiled list of properties, not even a term, or even existence, agreed upon; the official language is SMG.

With respect to child language acquisition, it should come as no surprise that to date no studies exist that investigate the nature, quality, and quantity of linguistic input children growing in Cyprus receive. There are simply no data available that would tell us about the proportion of basi- vs. acrolectal CG, purported CSG, and SMG in a young child's life, and whether there are differences between rural and urban upbringing or across different geographical locations. At this time, such information can only be estimated anecdotally.

We follow recent work from our research group, the Cyprus Acquisition Team (CAT), and adopt Rowe and Grohmann's (2013) term *(discrete) bilectalism* to characterize Greek Cypriot speakers in this diglossic speech community (for further discussion, see Grohmann and Leivada, 2012; Papadopoulou et al., 2014; Rowe and Grohmann, 2014), replacing our original notion of "bi-*x*" (Grohmann, 2011; Grohmann and Leivada, 2011). The first published study that addressed the role of bilectalism in language development, applied to lexical retrieval (Kambanaros and Grohmann, 2010, 2011; Kambanaros et al., 2010), is Kambanaros et al. (2013b), followed up by work comparing typically developing bilectal children to children with specific language impairment (Kambanaros et al., 2013a). To date, the lexical and morphosyntactic differences between CG and SMG qua bi-*x* or bilectalism have also featured in work on adult grammar (Grohmann and Papadopoulou, 2011) as well as specific topics in typical or impaired language, including light verb use (Grohmann and Leivada, 2013; Kambanaros and Grohmann, 2015), the comprehension and production of relative clauses (Theodorou and Grohmann, 2013), and the importance of creating an assessment tool for the diagnosis of specific language impairment for CG (Theodorou, 2013; Theodorou et al., submitted). We also raised the issue of bilectal populations as a topic of interest for research in bilingualism (e.g., Antoniou et al., 2014; Kambanaros et al., 2014), leaving the door open to classify these speakers as "bilingual" after all, once a better definition of language proximity in multilectal speakers is available beyond some notion of "second dialect (acquisition)" (cf. Siegel, 2010); this is part of our research agenda for comparative bilingualism.

With all this in place, we can assume that Greek Cypriots are typically sequential bilectal, first acquiring CG and then SMG (or something akin, such as CSG), where the onset of SMG may set in with exposure to Greek television, for example (clearly within the critical period) but most prominently with formal schooling (around first grade, possibly before, where the relation to the critical period is more blurred). What is more, due to the close relations between Cyprus and Greece (beyond language for historical, religious, political, and economic reasons), we are able to tap into two further interesting populations, all residing in Cyprus (Leivada et al., 2010): Hellenic Cypriot children, who are binational having one parent from Cyprus (Greek Cypriot) and one from Greece (Hellenic Greek), and Hellenic Greek children, with both parents hailing from Greece. Anecdotally, we could then say that binational Hellenic Cypriot children are presumably simultaneous bilectals (strong input in SMG and CG from birth), while Hellenic Greek children are arguably as close to monolingual Greek speakers in Cyprus as possible (SMG-only input from birth), though with considerable exposure to the local variety (CG)—again, certainly, once they start formal schooling.

Just as language development in bilingual children should be compared to that of monolinguals, different language combinations in bi- and multilingual children should be taken into consideration as well. Let us call this approach "comparative bilingualism," although in a very different conception from occasional mentions in the literature that deals largely with societal and educational issues in bilingualism (cf.

Bernbaum, 1979; Baker, 1996). In the next section, we will present our research group's findings on the acquisition and subsequent development of object clitic placement by bilectal and multilingual children in Cyprus. Looking at the four purported dynamic metrics of assessment, we may not yet know how much Greek *input* the bilingual children in Cyprus receive, and how SMG-like it is (which also holds for the bilectals). The same goes for the age of *onset* of SMG, if indeed prior to formal schooling, or the exact role of CSG in this respect. However, we do know for *timing* that object clitics appear very early in Greek (for SMG see Marinis, 2000, and for CG Petinou and Terzi, 2002, as well as our own CAT lab research reported below). And lastly, with respect to language *proximity*, CG as a "dialect" of Modern Greek is by definition very close to SMG (as opposed to, say, Russian). A valuable tool for further teasing apart timing and proximity from onset and input is Tsimpli's (2003) Interpretability Hypothesis (cf. Tsimpli and Mastropavlou, 2007), which has recently been assessed for Russian–Greek-speaking adults residing in Cyprus (Karpava, 2014), though we do not yet have comparable data from Russian-speaking bilingual children growing up in Greece (with SMG), which is part of our ongoing research activities: There does not seem to be a correlation between age/onset/input and the production of clitics, for example, which express uninterpretable features—and for which native-like attainment cannot be reached.

THE CAT CLITIC CORPUS: ACQUISITION AND DEVELOPMENT OF CLITIC PLACEMENT

One of the best studied grammatical differences between the two varieties pertains to clitic placement (see Agouraki, 1997, and a host of research since): Pronominal object clitics appear postverbally in CG, with a number of syntactic environments triggering proclisis, while SMG is a preverbal clitic placement language in which certain syntactic environments trigger enclisis. In both varieties of Modern Greek, 3rd person object clitics are derived from strong pronouns; clitics are marked for number (singular, plural), gender (masculine, feminine, neuter), and case (accusative, genitive). Concerning the particular characteristics of mixed clitic placement, it can be observed that certain syntactic environments enforce preverbal placement—otherwise enclisis is found. Therefore, clitics in CG can appear postverbally in both imperative and non-imperative contexts, whereas in SMG they can only appear as enclitics in imperatives and gerunds.

Now, the acquisition of pronominal clitics is arguably a "(very) early phenomenon," as Tsimpli (2014) calls it, since clitics represent a core aspect of grammar and are fully acquired at around 2 years of age. Using a sentence completion task that aimed at eliciting a verb with an object clitic in an indicative declarative clause (Varlokosta et al., 2015), we counted children's responses to the 12 target structures in CG, which should consist of verb–clitic sequences (as opposed to clitic–verb in SMG). Methodology and participant details will be provided below. To anticipate the presentation of results, the main pattern is

consistent with the one originally reported for our first pilot study (Grohmann, 2011), which was confirmed and extended to many more participants in subsequent work (summarized in Grohmann, 2014a). This pattern is provided in **Figure 1**.

With very high production rates in all groups (over 92%), the pilot study showed that the 24 three- and four-year-old children behaved like the 8 adult controls: 100% enclisis in the relevant context. In contrast, the group of 10 five-year-olds showed mixed placements, where that group is split further into three consistent sub-groups. The following introduces in some detail the CAT Clitic Corpus of data we have collected to date and briefly presents the main tool(s) used to elicit the responses (from Grohmann, 2014a). This level of detail also underlines our urge for more carefully controlled experimental investigations in the future (picked up in the Discussion and Outlook section). There are numerous references to our published works which each only consider smaller sub-groups; unfortunately, we cannot provide the overall analysis here, since it has not yet been published (Grohmann et al., submitted). For this reason, the presentation of the results below will be rather short and general, but the direction where this research project is heading and the relevance to the present contribution should become clear.

If we only consider the typically developing bilectal Greek Cypriot children that participated in the studies reported in Grohmann (2014a), we currently have 623 datasets of 12 target structures each; for these, we also have 34 adult controls and 20 teenagers, and we can compare them to additional populations, all residing in Cyprus: bilectal children with atypical language development (SLI), bilingual or rather bilectal bilingual children (Russian–Greek), Hellenic Cypriot or binational children (SMG and CG), and Hellenic Greek children and adults (SMG). These groups yield a total of 787 individuals that participated in the clitics tool(s). Most of these were Greek Cypriot children, but there are a number of other participants, as just listed. Likewise, most testing was done on the Clitics-in-Islands tool (COST Action A33, 2006–2010), presented below, but other tasks were used, too (see Grohmann, 2014a, for details and references). Here we focus on reporting data collected on the acquisition and subsequent development of object clitic placement in the two closely related varieties of Greek by bilectal (Grohmann, 2011, 2014a; Grohmann et al., 2012), binational (Leivada et al., 2010), and bilectal bilingual or multilingual children (Karpava and Grohmann, 2014).

FIGURE 1 | Clitic placement in clitics-in-islands task (all tested groups). From Grohmann (2014a, p. 196).
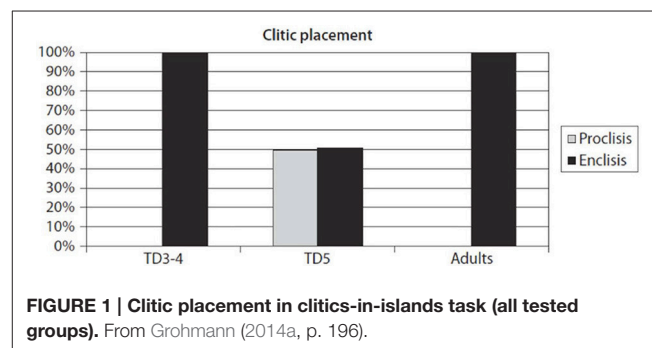
As shown in **Table 1**, this total number of participants breaks down as follows: 727 children from public kindergarten, pre-school, and primary school, 20 teenagers from public middle and high school, and 40 adults from university and the general employment sector, with an eye on gender balance. Of the 727 children, all but 34 had typical language development to the best of our knowledge. 623 were "monolingual" Greek Cypriot children (i.e., *sequential bilectal* in CG and SMG), 40 "monolingual" Hellenic Greek children (native in SMG but exposed to CG due to residence in Cyprus), and 30 binational Hellenic Cypriot children (native in SMG and CG, possibly with a preference for SMG from early on, but otherwise idealized *simultaneous bilectal*). In addition, 18 children were bilectal bilingual (Russian and Greek, i.e., CG and SMG), all with Russian-speaking mothers and Greek Cypriot fathers, but not tested for language delay or impairment, and the remaining 16 bilectal children were diagnosed with SLI by experienced speech–language therapists.

All participants from the studies reported in Grohmann (2011), Grohmann et al. (2012), and Theodorou and Grohmann (2015) were semi-randomly recruited across the urban centers of Nicosia and Limassol. The children from Leivada et al.'s (2010) study all came from the Nicosia municipality, and the bilingual children from Karpava and Grohmann (2014) all grew up in the Larnaca area. Due to the nature of the investigation (see Agathocleous et al., 2014), the children recruited for Agathocleous (2012) and Charalambous (2012) not only came from all over Cyprus (minus Nicosia and Limassol) but were also balanced for urban vs. rural upbringing. The reason for these details lies in the often raised but largely anecdotal claim that there is geographically based dialectal variation in Cyprus and that rural CG differs from urban CG. While this may be the case in many domains of the language (such as, most obviously, the lexicon), it did not seem to make a difference for the clitics task at hand, though in the absence of an empirically grounded knowledge base, we had to go to lengths to determine said absence of effects.

Further prerequisites for child participation included the following (with the exception of the Russian–Greek children from Karpava and Grohmann, 2014): Children had to attend Greek-speaking nurseries or kindergartens, be monolingual (i.e., bilectal) speakers of CG, and not have received speech–language therapy services. They were tested upon written parental consent and with approval from the Cyprus Ministry of Education and Culture (through the Pedagogical Institute). Of the older participants, 20 Greek Cypriot teenagers and 28 Greek Cypriot adults were tested who were all born and raised in Cyprus and resided in Cyprus at the time of testing; none of the teenagers had spent any large amounts of time outside the country. In addition, 6 Hellenic Greek adults residing in Cyprus were tested. None of the older participants was reported to have had speech, language, or communication difficulties.

In sum, what this line of research focuses on is a comparable "linguality" of participants, here children that grow up with one language (Greek) which comes in at least two distinct (i.e., discrete) lects, CG and SMG, leaving aside the issue of CSG. The attribute of linguality goes beyond, or in addition to,

TABLE 1 | Breakdown of all participants (clitic tasks).

Participant groups (ethnicity/language)	Number	Age range	Gender
Bilectal children (Greek Cypriot/CG and SMG)	6	2;8–2;11	331F, 292M
	23	3;0–3;11	
	154	4;0–4;11	
	193	5;0–5;11	
	185	6;0–6;11	
	36	7;1–7;11	
	26	8;1–8;11	
Monolingual children (Hellenic Greek/SMG)	2	3;2	23F, 17M
	10	4;0–4;10	
	8	5;1–5;11	
	1	6;3	
	12	7;0–7;11	
	7	8;1–8;10	
Binational children (Hellenic Cypriot/SMG and CG)	2	3;7	22F, 8M
	2	4;1–4;2	
	8	5;2–5;10	
	7	6;0–6;11	
	5	7;0–7;10	
	5	8;0–8;7	
	1	9;1	
Bilingual children (Russian–Cypriot/R, CG, SMG)	2	4;8	7F, 11M
	2	5;4–5;6	
	9	6;0–6;8	
	5	7;0–7;8	
Bilectal children with SLI (Greek Cypriot/CG and SMG)	1	4;11	6F, 10M
	8	5;3–5;11	
	1	6;7	
	3	7;1–7;10	
	3	8;1–8;7	
TOTAL (CHILDREN)	727	2;8–9;1	389F, 338M
Bilectal teenagers (Greek Cypriot/CG and SMG)	20	14–18 (mean: 15;6)	11F, 9M
Bilectal adults (Greek Cypriot/CG and SMG)	34	20–65 (mean: 38;6)	17F, 17M
Monolingual adults (Hellenic Greek/SMG)	6	20–30 (mean: 23;6)	2F, 4M

CG, Cypriot Greek; F, Female; M, Male; SMG, Standard Modern Greek (from Grohmann, 2014a, p. 11).

whether a child may also grow up bilingually (simultaneously or sequentially) or learn additional languages later on. In the absence of (i) relevant studies concerning quality and quantity of lectal input, age of onset, and other important factors for the early years, as well as (ii) a clear characterization of acrolectal CG as CSG and (iii) its relevance for child language development, we have to leave things here as they stand and idealize somewhat. It is in this sense that we describe the linguality of Greek Cypriots as (discrete) bilectalism.

For the purpose of this research, the COST Action A33 Clitics-in-Islands testing tool (Varlokosta et al., 2015)—originally

designed to elicit clitic production even in languages that allow object drop, such as European Portuguese (Costa and Lobo, 2007)—was adapted to CG (from Grohmann, 2011). This tool is a production task for a 3rd person singular accusative object clitic within a syntactic island in each target structure in which the target-elicited clitic was embedded within a *because*-clause (where the expected child response is provided in brackets and the clitic boldfaced):

(1) To aɣori vreʃi ti   ɣata tʃe i   ɣata e vremeni. Jati i   ɣata e vremeni?
     the boy wets the cat and the cat is wet   why the cat is wet

     I   ɣata e vremeni jati   to aɣori… [vreʃi   **tin**].
     the cat is wet   because the boy   wet.PRES.3SG   CL.ACC.3SG.FEM

     'The boy is spraying the cat and the cat is wet. Why is the cat so wet? The cat is wet because the boy… [is spraying it].'

The task involved a total of 19 items; 12 target structures (i.e., test items) after 2 warm-ups, plus 5 fillers. All target structures were indicative declarative clauses formed around a transitive verb, with half of them in present tense and the other half in past tense. Children were shown a colored sketch picture on a laptop screen, depicting the situation described by the experimenter. The scene depicted in **Figure 2** corresponds to the story and sentence completion in (1), for example.

Other test examples can be found in Agathocleous et al. (2014), who also discuss the "short version" in some detail (a pre-version developed within COST Action IS0804, 2009–2013), as well as Karpava and Grohmann (2014), who in addition present the Production Probe for Pronoun Clitics tool (based on Tuller et al., 2011).

Combining the different tasks and participant details, our growing CAT Clitics Corpus—and as yet not fully statistically analyzed beyond what is reported here (though for a first attempt see Grohmann et al., submitted)—at present contains data from a host of participants (Grohmann, 2014a, p. 14). These details are summarized in **Table 2**, where the boldfaced row indicates the total numbers of participants tested on a comparable tool, namely some version of the above-described elicitation tool for CG with 12 identical target-elicitation structures in either version.
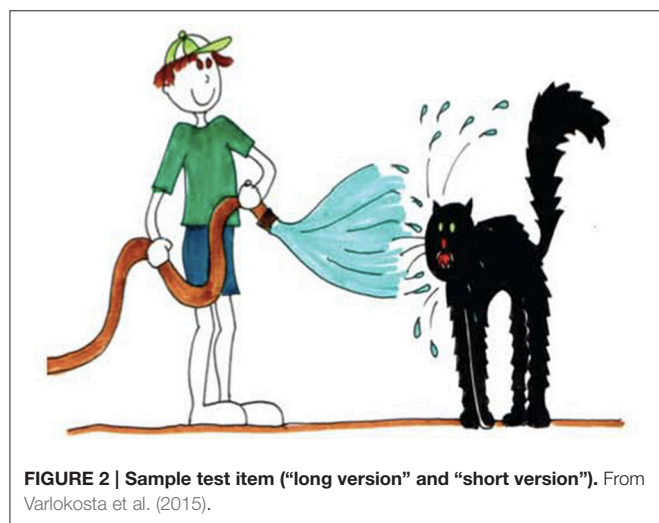
All tests with Greek Cypriot bilectal children were carried out by native speakers of CG; those tests that were administered in SMG were done by a native SMG speaker. Testing was conducted in a quiet room individually (child and researcher). Most children were tested in their schools or in speech–language therapy clinics, but a few were tested at their homes. It is well known that Greek Cypriots tend to code-switch to SMG or some hyper-corrected form of "high CG" when talking to strangers or in formal contexts, as mentioned by Arvaniti (2010), Rowe and Grohmann (2013), and references cited there. For this reason, in an attempt to avoid a formal setting as much as possible (and thus obtain some kind of familiarity between experimenter and child), a brief conversation about a familiar topic took place before the testing started, such as the child's favorite cartoons.

All participants received the task in one session, some in combination with other tasks (such as those tested in Theodorou and Grohmann, 2015; see Theodorou, 2013). The particular task lasted no longer than 10 min, the "short version" even less. The pictures were displayed on a laptop screen which both experimenter and participant could see. The child participants heard the description of each picture that the experimenter provided and then had to complete the *because*-clause in which the use of a clitic was expected; some participants started with *because* on their own, others filled in right after the experimenter's prompt of *because*, and yet others completed the sentence after the experimenter continued with the subject [the bracketed part in (1) above].

No verbal reinforcement was provided other than encouragement with head nods and fillers. Self-correction was not registered; only the first response was recorded and used

FIGURE 2 | Sample test item ("long version" and "short version"). From Varlokosta et al. (2015).

TABLE 2 | Number of participants (per tool).

Clitic elicitation tool	Bilectal children	Other children	Teenagers and adults
Clitics-in-Islands tool ("long version": CG)	443 16 (SLI)	18 (BL) 30 (BN) 40 (HG)	10 (GC T) 24 (GC A) 6 (HG A)
Clitics-in-Islands tool ("short version": CG)	180	—	10 (GC T) 10 (GC A)
TOTAL	623 16 (SLI)	18 (BL) 30 (BN) 40 (HG)	20 (GC T) 34 (GC A) 6 (HG A)
Clitics-in-Islands tool ("modified long": SMG)	40	30 (BN) 40 (HG)	6 (GC A) 6 (HG A)
Production Probe for Pronoun Clitics tool (CG)	—	18 (BL)	—

A, adult; BL, bilingual; BN, binational; CG, Cypriot Greek; GC, Greek Cypriot; HG, Hellenic Greek; SLI, specific language impairment; SMG, Standard Modern Greek; T, teenager. From Grohmann (2014a, p. 15).

for data collection and analysis purposes. Regardless of the child's full response, all that was counted were verb–clitic sequences (for clitic production) and the position of the clitic with respect to the verb (for clitic placement). Except for the studies reported in Agathocleous et al. (2014), the experiments were not audio- or video-taped, but answers were recorded by the researcher or the researcher's assistant on a score sheet during the session; many testing sessions involved two student researchers with one carrying out the task and the other recording the responses (in alternating order). In those studies in which different clitic tasks were administered (Karpava and Grohmann, 2014), or where the same tool was tested in CG and SMG (Leivada et al., 2010), participants were tested with at least 1 week interval in between.

(DISCRETE) BILECTALISM AND THE SOCIO-SYNTAX OF LANGUAGE DEVELOPMENT

All these different studies with different populations and different age groups but the same tool show the following. First, the production rate of clitics in this task is very high from an early age on, safely around the 90% mark from the tested age of 2;8 onwards (lowest production at around 75%), over 95% at age 4;6 (lowest production at around 88%), and close to ceiling for 5-year-olds and beyond. The sub-group of 117 children from Grohmann et al. (2012) performed as shown in **Table 3**.

This said, Leivada et al. (2010) found considerably higher productions for the younger Hellenic Greek and Hellenic Cypriot children tested compared to their Greek Cypriot peers. However, just considering the 623 bilectal children, we can confirm that the task was understood and elicited responses appropriate; in the widely tested age group of 5-year-olds, the production numbers are among the highest of all languages tested (Varlokosta et al., 2015). High production means reliable data points for all 12 target structures; statistical analysis confirms that there were neither item effects nor test effects, that is, the productions for the "long" and "short" version of the clitics tool are fully comparable (Grohmann, 2014a).

Second, and most importantly, the analysis of the 431 datasets of the bilectal children presented by Grohmann et al. (submitted) are consistent with the findings of the much smaller pilot study (Grohmann, 2011). In other words, **Figure 1** can be used as a general indicator: Up to around age 4, children reliably produce

enclisis in this task at just shy of 90%, as expected (and confirmed by adult speakers), while we find considerable variation in clitic placement in the 5- to 7-year-olds.

To illustrate with the subset of 117 children again, when their non-target preverbal clitic placement productions were plotted according to chronological age, the resulting curve looks as in **Figure 3**.

However, what we can observe are apparent inconsistencies in terms of clitic placement, in particular by comparing younger with older children according to their schooling level. While for nursery children (mean age 3;3), target postverbal clitic placement lies at 93%, it decreases systematically for each additional year of formal schooling: kindergarten (4;3) at 82%, pre-school (5;5) at 73%, and first-grade (6;7) at 47%—from grade 2 onwards, the rates quickly shoot up toward 100% again (Grohmann, 2014a). This analysis is extended in Grohmann et al. (submitted). But using the same sub-group of 117 children again, compare **Figure 3** above with **Figure 4**.

The most striking result is that, while at the youngest ages, prior to formal schooling, the CG-target enclisis is produced predominantly, if not exclusively, once Greek Cypriot children start getting instructed in the standard language (SMG or some such equivalent like CSG), their non-target productions of proclisis rise dramatically—all the way to second grade (not shown here; full analysis provided in Grohmann et al., submitted).

One obvious way to approach the situation is to appeal to "competing grammars." Kroch (1994: 180) proposes competition of grammatical systems for diachronic change in that "syntactic change proceeds via competition between grammatically incompatible options which substitute for one another in usage" (for specific accounts and extensions to language acquisition models, e.g., Kroch and Taylor, 2000; Yang, 2000; Legate and Yang, 2007). Following Lightfoot's (1999) description of competing grammars reflecting "internalized diglossia," this might indeed be a good approach to take up for CG. In fact, Tsiplakou (2009, 2014) had already addressed a possible implementation of the competing-grammars hypothesis for

TABLE 3 | Clitic production (adapted from Grohmann et al., 2012).

Age range (Number)	Overall clitic production (%)	Target postverbal clitic placement (%)
2;8–3;11 (N = 26)	89.4	89.2
4;0–4;11 (N = 21)	88.5	88.0
5;0–5;11 (N = 50)	94.3	68.0
6;0–6;11 (N = 20)	87.3	47.0
adult controls (N = 8)	100	100

From Grohmann (2014a, p. 17).
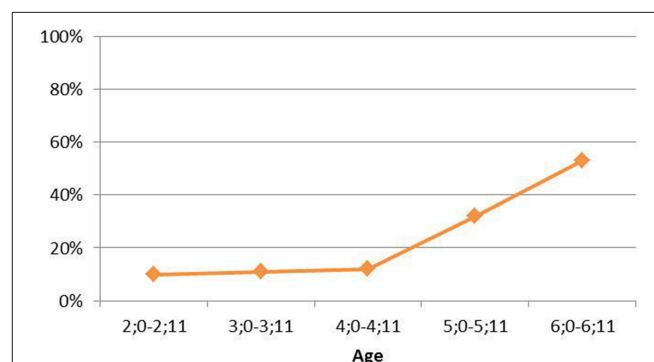
FIGURE 3 | Non-target preverbal clitic placement (by chronological age). The x-axis indicates participants according to their chronological age, while the y-axis plots non-target preverbal clitic placement in the participants' responses (percentage). From Grohmann and Leivada (2011).

FIGURE 4 | Non-target preverbal clitic placement (by schooling level).
The x-axis indicates participants according to their schooling level, while the
y-axis plots non-target preverbal clitic placement in the participants' responses
(percentage). From Grohmann and Leivada (2011).
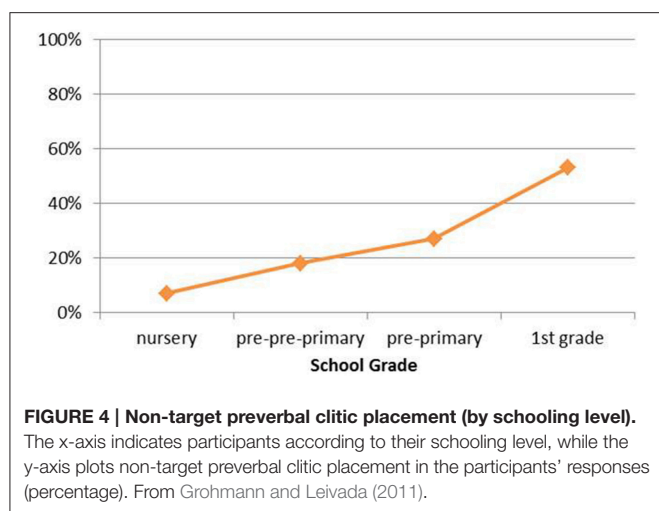
CG; for further discussion, as well as the extension to the
older notions of "competing motivations" (Du Bois, 1985) and
"metalinguistic awareness" (Cazden, 1976, see Leivada and
Grohmann, 2016).

Such an approach would pit the native CG grammar (in
this case: enclisis) against the emerging SMG grammar (here:
proclisis), which happens to grow stronger through increased
input. Since formal schooling is carried out, by law, in the
medium of SMG, it is around the entrance into the public
schooling system that the SMG grammar becomes stronger,
perhaps even dominant at times. This does not imply, however,
that public schools in Cyprus would constitute a monolingual,
monodialectal environment for pupils. Classroom studies have
shown that "CG is very often used as a medium of interaction and
even instruction during classroom," as a reviewer reminded us,
across all grades (e.g., Yiakoumetti, 2007; Sophocleous and Wilks,
2010; Sophocleous, 2011).

We would like to take these findings one step further and
suggest that they are best captured by the *Socio-Syntax of
Development Hypothesis* (Grohmann, 2011), namely that an
explicit "schooling factor" is involved in the development of the
children's grammar. Note that this grammatical development
takes place past the critical period and does so possibly in
combination with "competing motivations" (Grohmann and
Leivada, 2011; Leivada and Grohmann, 2016). These arguably
stem from the (at least) two grammars in the bilectal child's
linguistic development that compete with each other. In other
words, the Socio-Syntax of Development Hypothesis can be
seen as the specific trigger for competing grammars in the
development of CG clitic placement by young children.

A way to appreciate the more general Socio-Syntax of
Development Hypothesis would be to approach the acquisition
of syntactic variants, which CG enclisis and SMG proclisis in
the same environment arguably are, by assuming competing
motivations that arise between the home and the school variety.
In the present case, CG enclisis competes with SMG proclisis
in the same syntactic context between two varieties in a
dialectal continuum which thus have close proximity. Given

that all schooling is done through the medium of SMG, the
relevant competing motivations in Cyprus may derive from
the absence of bilectal education that could increase children's
awareness of the low social prestige of their native CG (see
also Rowe and Grohmann, 2013, for further discussion and
references).

Note that the rate of 100% proclisis in the Hellenic Greek
children is by no means an accidence. A study carrying out
the identical tool in Greece (Varlokosta et al., 2014) found
that children aged between 3;6 and 5;11 as well as children
with SLI exclusively produced proclitic placement of the direct
object clitics—as was expected, since SMG does not allow for
enclisis in the environment tested (also reported in Varlokosta
et al., 2015). A similar point can be made for the binational
Hellenic Cypriot children, who performed more like the Hellenic
Greek children (in Greece and Cyprus) than their Greek Cypriot
bilectal peers. Here we might find a possible difference in
development for simultaneous vs. sequential bilectals: If on the
right track, Hellenic Cypriot children, having simultaneously
acquired CG and SMG, do not enter into competition due to
confusion or increased SMG input; both varieties are perfectly
natural sources of linguistic input from birth. In addition, as
fully balanced users of both, they do not enter competing
motivations either but are already metalinguistically aware of the
two systems and their appropriate use. (See also the next section
for added evidence coming from cognitive abilities, though
Hellenic Cypriot children need yet to be assessed, which is part
of an ongoing dissertation under the first author's supervision).

Lastly, we also collected data from a group of clear-cut bi-
or multilingual children in Cyprus: Russian–Greek speakers,
particularly those with a Russian-speaking mother and Greek
Cypriot father, whose languages are thus Russian, CG, and
SMG (Karpava and Grohmann, 2014); in fact, these children
are perhaps best labeled "bilectal bilingual." Comparing our data
from 18 bilectal bilingual children on the same tool with 40
bilectal children (Leivada et al., 2010), we note the following
stark contrast in target postverbal clitic placement (with almost
identical production levels): kindergarteners at only 22% enclisis
(SD 2.08, compared to 82% for bilectals), pre-schoolers at 8%
(SD.71, compared to 73%), and first-graders at 11% (SD 3.26,
compared to 47%).

Clitic placement thus shows that the bilingual children
increased their usage of proclisis and decreased enclisis from
kindergarten to primary school. In contrast to the bilectal
children, they exhibited much more proclisis than target enclisis
early on. This may be due to the additional presence of SMG in
the family environment rather than CG-only: Due to L2 learning
through formal instruction, most of the Russian mothers' input
when addressing their child in Greek (which is quite frequent)
would be more SMG-like. In addition, they tend to have a
negative attitude toward CG. Since the bilingual children also
have higher metalinguistic awareness, they are influenced by their
mothers as well as their peers: The former often exhibit a negative
attitude toward the CG variety, while the latter arguably show
a strong preference toward it. At school, they are forced to use
SMG, which is in line with their mothers' linguistic behavior, but
contrasts with their peers' and their fathers'. In this sense, they are

constantly urged to not only make a choice of language (Russian vs. Greek), but also of variety (CG vs. SMG), and this choice seems to be influenced by different factors.

Let us phrase this in the context of Tsimpli (2014). While clitic acquisition in terms of production is not a problem for simultaneous bilingual children, the appropriate use is somewhat more tricky. First, it is known that non-core aspects of language license the appropriate use and interpretation of clitics, such as pragmatics and discourse/context sensitivity. These are particularly relevant for bilingual populations who acquire a clitic language alongside a non-clitic language (such as Russian), for which the appropriate referent choice is often at stake (full DP vs. strong vs. clitic pronoun), as Parodi and Tsimpli (2005), among others, have shown. In addition, we are dealing with a different situation which lies clearly outside core grammar: the sociolinguistically appropriate placement of clitics. We observe that both bilectal and bilingual children struggle with the context-appropriate form, which arguably involves a certain amount of maturation and metalinguistic awareness.

A GRADIENCE OF THE COGNITIVE ADVANTAGE OF BILINGUALISM?

We will now turn to a first study on the purported bilingual status of Greek Cypriot bilectal children and its relevance for a more gradient, comparative bilingualism. The results from a range of executive control (EC) tasks administered to monolingual SMG-speaking children (in Greece) as well as CG–SMG bilectal and Greek–English bi-/multilingual children (in Cyprus) suggest that bilectal children behave more like their multilingual rather than their monolingual peers (Antoniou et al., 2014)—that is, on a scale in between.

It has frequently been suggested that bilingualism bears an impact on children's linguistic and cognitive abilities (see recent overviews by and the literature cited in Kroll and Bialystok, 2013; Barac et al., 2014). For example, as already mentioned above in the context of Tsimpli (2014), bilingual children arguably have smaller vocabularies in each of their spoken languages as a result of input deficit (e.g., Paradis and Genesee, 1996; Oller and Eilers, 2002; Unsworth, 2013). On the other hand, bilingual children seem to exhibit earlier development of pragmatic abilities: They are more advanced in computing scalar implicatures (Siegal et al., 2007) and better in detecting violations of Gricean maxims (Siegal et al., 2009, 2010), for example; bilingual children presumably compensate for their lower lexical knowledge by paying more attention to contextual information. And then there is the long-standing claim that bilingualism enhances children's development of EC, the set of cognitive processes that underlie flexible and goal-directed behavior, commonly referred to as the "bilingual advantage" or "cognitive advantage of bilingualism" (for overviews, e.g., Bialystok, 2009; Baum and Titone, 2014; Costa and Sebastián-Gallés, 2014; see also the meta-analysis provided by Adesope et al., 2010). Taking a particular influential one of the many approaches to EC, there is a tripartite distinction into working memory, task-switching, and inhibition (Miyake et al., 2000),

each with their own rationale, though more recently some doubt has been cast on inhibition as a separate executive component (Miyake and Friedman, 2012).

Starting with the latter, a bilingual advantage in inhibition may relate to the ability to suppress dominant, automatic responses or irrelevant information (e.g., de Abreu et al., 2012; Poarch and van Hell, 2012). There is also some evidence for advanced task-switching, that is, the ability to flexibly switch attention between rules (e.g., Bialystok and Viswanathan, 2009; Foy and Mann, 2014). The effect of bilingualism on working memory, the ability to simultaneously maintain and manipulate information in mind, is more controversial, however (e.g., de Abreu, 2011; Morales et al., 2013; Blom et al., 2014; Calvo and Bialystok, 2014).

This composite approach to EC is arguably superior to an earlier suggestion that the bilingual advantage can be traced exclusively to more advanced inhibition alone (e.g., Bialystok, 2001). Here the idea was that, because both linguistic systems are activated when a bilingual speaks in one language, fluent use requires the inhibition of the other language. This constant experience in managing two active conflicting linguistic systems via inhibition enhances bilinguals' inhibitory control mechanisms. This early view, however, has been challenged on several grounds (e.g., Bialystok et al., 2012). One line of argument would be that advantageous effects of bilingualism have been observed for the very first years of life, even for 7-month-old infants (Kovács and Mehler, 2009). Since language production has not yet started in bilingual infants, there would be no need to suppress a non-target language. We are not sure that this argument goes through, though: After all, even bilingual infants are fully aware of the different languages they are acquiring, and while they may not need to inhibit one to produce the other, they presumably process the two (or more) languages and should therefore regularly inhibit one to process the other. However, there are a number of further arguments to take a more differentiated view on EC as the measuring stick for the bilingual advantage, as put forth in many of the references cited above; see also Antoniou (2014) and Antoniou et al. (2016) for further discussion.

All in all, an advantage in EC may be the result of constantly having to manage two different linguistic systems. One aspect of continued research on the topic would thus be to disentangle the different EC sub-components and determine which aspect(s) of executive control really relates to a bilingual advantage. Regarding performance on executive control in monolingual, bilectal, and bi- or multilingual children, the relevant research question is then (Antoniou et al., 2014): What is the effect of bilectalism on children's vocabulary, pragmatic, and EC skills?

A total of 136 children with a mean age of just above 7.5 years of age participated in the study (Antoniou et al., 2014): 64 Greek Cypriots, bilectal in CG and SMG, aged 4;5–12;2 (mean age: 7;7, SD: 1;6 years; 32 boys, 32 girls); 47 residents of Cyprus, multilingual in CG, SMG, and English (plus in some cases an additional language), aged 5;0–11;5 (mean 7;8, SD 1;8; 24 boys, 23 girls); and 25 Hellenic Greeks, monolingual speakers of SMG, aged 6;2–9;0 (mean 7;4, SD 0;9; 15 boys, 10 girls). Socio-economic status measures included the Family Affluence Scale (Currie et al., 1997), while level of maternal and paternal education, among

other details, were obtained through questionnaires (Paradis et al., 2010; Paradis, 2011). Since the multilingual children all attended a private English-medium school, their socio-economic was higher than the mean of all other participants.

A range of language proficiency measures were administered for expressive and receptive vocabulary, including the Greek versions of the Word Finding Vocabulary Test for expressive vocabulary and the revised Peabody Picture Vocabulary Test (SMG) as well as the Greek Comprehension Test (for either variety). For pragmatic performance, a total of six tools were used, tapping into relevance, manner implicatures, metaphors, and scalar implicatures; the bilectal and multilingual children received the test in CG, the bilectals took the test in both CG and SMG, and the monolinguals were tested in SMG only. As for non-linguistic performance, the WASI Matrix Reasoning Test was used to assess participants' non-verbal intelligence. The EC tasks administered included a wide range of batteries. For verbal working memory, the Backward Digit Span Task was employed, and for visuo-spatial working memory, an online version of the Corsi Blocks Task. Inhibition was assessed through Stop-Signal and the Simon Task, and switching through the Color–Shape Task. (For more details and references, see Antoniou, 2014; Antoniou et al., 2014.) In the end, we opted for a composite measure of EC which was computed in a principled component analysis for the factors Working Memory and Inhibition over the individual results (Antoniou et al., 2016).

The analysis results from a two-stage comparisons between the three groups. First, the performance of all child participant groups was compared to each other (monolinguals vs. bilectals vs. multilinguals); the three groups were matched in age by excluding all children who were below 6 and above 9 years of age. Then the performance of a subset of 17 bilectal children was compared to that of the monolingual group. All these children were also administered a receptive vocabulary test in order to test whether exercising a more rigid statistical control over children's language skills would reveal or increase potential bilectal advantages in EC. As Antoniou et al. (2016) show, the two composite measures (Working Memory and Inhibition) significantly and positively correlate with language ability; also, the bilectal children were possibly disadvantaged in language proficiency relative to monolinguals.

The results from this study can be presented across four types of group comparisons. The first concerns background measures. The relevant subsets of the three participant groups of bilectal ($n = 44$), multilingual ($n = 26$), and monolingual children ($n = 25$) aged 6;0–8;11 were intended to be matched for age and gender; they did not statistically differ on age [$F_{(2, 92)} = 0.696$, $p > 0.05$] or gender [$F_{(2, 92)} = 0.587$, $p > 0.05$]. However, they did differ on socio-economic status [$F_{(2, 89)} = 9.622$, $p < 0.05$], with the private-schooled multilingual children as a group coming from a higher socio-economic family background than the monolingual ones, and the bilectals from the lowest. The three groups also differed on non-verbal IQ [$F_{(2, 92)} = 3.377$, $p < 0.05$], with the multilingual children higher than the two other groups, which did not differ significantly.

Next we compared the three participant groups' performance on the vocabulary measures. The multilingual children had a significantly lower vocabulary score than the bilectals, who in turn had a significantly lower vocabulary than the monolinguals [$F_{(2, 92)} = 44.183$, $p < 0.05$], confirmed by *post-hoc* pairwise comparisons with Bonferroni correction for multiple comparisons (all ps < 0.05). From what is known about vocabulary growth in bilingual contexts (see references above), it was expected that the monolingual children would outperform the multilinguals; the fact that the bilectals fall in between fits nicely with our hypothesis that, on a gradient scale, bilectalism lies somewhere in between mono- and multilingualism.

The third group comparison concerns performance in the pragmatic tasks (Antoniou et al., 2014; this is not part of the extended analysis reported in Antoniou et al., 2016). Analyses of covariance (ANCOVAs), with vocabulary and SES & IQ as covariates, showed no significant differences between the three groups across all pragmatics tasks [$F_{(2, 87)} = 4.081$, $p < 0.05$]. No differences in the pragmatic tasks suggest that even those children who exhibit some sort of lower language (multilinguals, perhaps bilectals), they still show comparable pragmatic performance at the same age. With an eye on the Greek Cypriot bilectal children, this again suggests that they pattern somewhere in between; given the lower vocabulary scores compared to their monolingual peers from Greece, they do perform the same in the six pragmatic tasks.

Lastly, and for the purposes of our research question perhaps most importantly, the child participants' performance on the EC tasks was analyzed and submitted to principal component analysis (Antoniou et al., 2014). All three global EC scores (working memory, inhibition, and switching) positively correlated with IQ. ANCOVAs on the three composite scores for EC, with Group as a between-subjects factor and IQ, linguistic knowledge (Greek), age, and SES as covariates, revealed a significant effect of group only for the overall EC score: a significant multilingual advantage over monolinguals, with a trend for a bilectal advantage.

We illustrate this finding here with switch cost from the original analysis (Antoniou et al., 2014): Bilectals performed better than monolinguals in the congruent switch trials, with no other significant comparisons [$F_{(2, 87)} = 4.081$, $p < 0.05$]; in the incongruent switch trials, bilectals also performed better than monolinguals [$F_{(2, 87)} = 5.805$, $p < 0.05$], with multilinguals almost better than monolinguals ($p = 0.108$). These results can be summarized as showing that the bilectal children performed better than the monolinguals in overall EC ability and slightly worse than multilinguals. With respect to the lack of a clear effect for switching, as opposed to vocabulary, for example, we would like to suggest that there is an interference from language proximity: The more similar the two varieties, the more difficult it is to switch—or rather, the less there is a need to switch. For example, in a given group of individuals of whom all but one speak Greek and English, with one knowing no Greek, a Greek-language discussion would be translated or summarized in English for that individual [switching by the bilingual speaker(s)]. In contrast, in a group of Greek speakers of whom only one does not speak Cypriot Greek, a CG-at large discussion would arguably not be translated or summarized in SMG for that individual [no switching by the

bilectal speaker(s)]. As noted in a different context by Runnqvist et al. (2012), this may in fact tie in with the reverse of a bilingual advantage, what they call the "bilingual disadvantage." Beyond the cases they examine (e.g., Ivanova and Costa, 2008; Costa et al., 2009), it has also been suggested that the cognitive advantage only surfaces in bilingual individuals who actually switch between their languages frequently (Prior and Gollan, 2011).

In the extended statistical analysis of Antoniou et al. (2016), it could furthermore be shown through a mixed ANCOVA that, while the Working Memory and Inhibition composite scores significantly correlated with IQ, general language ability in Greek, and age, multilinguals had a significantly higher EC performance than monolinguals ($p < 0.05$), without any significant differences between the other groups (all $ps > 0.05$, Bonferroni correction applied). Also, since the Group × EC interaction was not significant [$F_{(2, 84)} = 0.744$, $p > 0.05$], the multilingual advantage in EC was not specific to Working Memory or Inhibition. Moreover, the second stage of the statistical analysis explores the possibility that a bilectal advantage over monolinguals can indeed be found if children's language proficiency in Greek is more rigidly controlled (see Antoniou et al., 2016).

In terms of a larger discussion, we hasten to add that there is recent work that casts some doubt on the purported relation between bilingualism and EC abilities (e.g., Paap and Greenberg, 2013; Paap and Sawi, 2014). Just like the above-mentioned modifications to the "right" kind of model of EC, there are a number of factors that make more careful investigations even more important. In the study reported here (Antoniou et al., 2014, 2016), for example, we compared group performances. However, the groups were composed of rather few children of a considerable age range, and, for obvious reasons for the populations chosen, there were significant differences in socio-economic status and non-verbal intelligence. Likewise, it is not yet clear in how much, if at all, the cognitive advantage observed in bilingualism pertains or increases in multilingualism. These are some of the considerations that our future work aims to improve in order to assess the purported bilingual advantage in EC abilities in bilectal speakers as well as finer grained and better selected multilingual groups for comparison.

An associated extension of the "bilingual advantage" in cognitive development for closely related varieties concerns children's development of literacy skills. This issue has recently been addressed for the two Norwegian literary varieties, Nynorsk and Bokmål, by Vangsnes et al. (2015). Although not directly linked to EC abilities, there is a growing body of work on literary development in Cyprus (Tsiplakou, 2006; Hadjioannou et al., 2011), but more recent research from Greece for SMG connects EC abilities explicitly with literary skills for mono- and bilingual children (Andreou, 2015; Andreou and Tsimpli, submitted). This connection is currently being investigated for bilectal, bilingual, and monolingual children at CAT as part of an ongoing dissertation under the first author's supervision.

LESSONS FROM DEVELOPMENTAL LANGUAGE IMPAIRMENT

In this third line of research related to the role of bilectalism within a comparative view to bi- or even multilingualism, we shift to studies that focus on the manifestations of lexical retrieval or spoken naming breakdown in atypical and impaired language development. The data reported come from our growing CAT Naming Corpus which includes monolingual, bilectal, bilingual, and multilingual child speakers of Greek. Here we aim to highlight the relevance of this research for a more gradient, comparative perspective of bilingualism in the context of developmental language impairments.

Lexical retrieval deficits, or childhood anomia, are a frequent part of the symptom complex that characterizes children with language impairments and are usually defined as "delayed or inaccurate responses with a high incidence of repetitions, reformulations, word substitutions, insertions, time fillers, and empty words" (German and Newman, 2004, p. 624). Speech and language therapists working with language-impaired children with anomia report co-existing impairments in other linguistic (expressive language, phonology, literacy) and non-linguistic domains (e.g., working memory); recent up-to-date reviews can be found in Friedmann et al. (2013) and Kambanaros et al. (2015). Depending on the severity, anomia may have severe repercussions for children in school settings, the significance being that classroom communication and academic skills, including reading and writing, are usually adversely affected (see Messer and Dockrell, 2011). Moreover, when anomia impedes communication with peers and others, children's psycho-social well-being is shown to be compromised (Tomblin, 2008). The emphasis lies on difficulties with lexical retrieval that manifests as an inability to name things like concrete entities (named by nouns) and actions (named by verbs).

We report on a study where the performance of multilingual children with SLI residing in Cyprus was compared with the performance of a language-matched group of multilingual children without SLI and with bilectal children, with and without SLI, on the same task. Multilingual children are in this context defined as children who simultaneously acquire two first native languages (e.g., CG and English) and SMG as a third language upon entering the school system, usually by the age of four (hence possibly falling under early second language acquisition; see Meisel, 2007); alternatively, one might refer to them as "bilectal bilinguals." The task used was a picture-based naming test of concrete noun and verbs, the Cypriot Object and Action Test (COAT; Kambanaros et al., 2013b). For a subgroup of the multilingual children with SLI, performance on noun and verb naming was investigated in two spoken languages (namely, Greek–English), using the English version of the OAT (Kambanaros, 2003, 2013).

A total of 59 children participated in the noun–verb naming study, divided into four groups:

- *bilSLI* ($n = 14$): 14 bilectal children with SLI (4 girls and 10 boys), aged 5;5–9;9 (average age 6;9, standard deviation 1;8)

- *multiSLI* ($n = 5$): 5 multilingual children with SLI (2 girls and 3 boys), aged 6;6–9;2 (mean age 7;11, standard deviation 1;1)
- *bilTLD-LM* ($n = 30$): 30 typically developing bilectal first-graders (15 girls and 15 boys), aged 6;0–6;11 (mean age 6;3, standard deviation 0;3), serving as the language-matched bilectal group for multilingual children with SLI
- *multiTLD-LM* ($n = 10$): 10 typically developing multilingual children (7 girls and 3 boys), aged 4;6–6;11 (mean age 5;1, standard deviation 0;9), serving as the language-matched control group for the multilingual children with SLI.

The children with typical language development (both, bilTLD-LM and multiTLD-LM) were recruited randomly from three public primary schools in urban Cyprus after approval from the Ministry of Education and Culture, and upon written parental consent. No typically language-developing child was or had ever been receiving speech–language therapy services. The children with SLI (both, bilSLI and multiSLI) were recruited from speech and language therapists in public primary education and/or private practice. All language-impaired children were in mainstream education and in the school grade corresponding to their chronological age. Also, they had received or were receiving speech–language therapy and/or special education services separate from their classmates and the regular classroom ("pull-in/out service model").

Subject selection criteria included: no history of neurological, emotional, or behavioral problems, no gross motor difficulties, hearing and vision adequate for test purposes, normal articulation, normal performance on screening measures of non-verbal intelligence (a score no less than 80 on the Raven's Colored Progressive Matrices or as reported by the school psychologist). All children came from families with medium to high socio-economic status. The bilectal children (both, bilTLD-LM and bilSLI) came from a Greek Cypriot background, with exposure to CG as the exclusive home language and SMG as the language of schooling. For the multilingual children (both, multiTLD-LM and multiSLI), a thus-defined bilectal background was required plus early exposure to a third non-Greek language in the home (such as English); in addition, all language acquisition involved bona fide multilingualism (e.g., a child exposed to CG and English from birth and later to SMG at school).

Of the five simultaneous multilingual children with SLI tested, three came from a CG–English language background, one was a CG–Romanian multilingual, and the other CG–Arabic. According to parental reports, all five multiSLI children were Greek-dominant. The group of multiTLD-LM, the typically developing multilingual preschoolers serving as the language-matched control group to the multilingual SLI group, were simultaneous bilinguals of CG (L1a) and a second language (L1b)—here: English, Romanian, Russian, and Arabic—and had acquired SMG as their L2 upon school entry (e.g., kindergarten at 4 years of age). In all cases, the father was of Greek Cypriot background and the mother a native speaker of the non-Greek language just specified. For all participating multilingual children, the Developmental and Language Background questionnaire developed in COST Action IS0804 (2009–2013), which both authors participated in, was given to the mothers to complete (see Tuller, 2015). Further information can be obtained from the authors.

Participating bilectal SLI or bilectal language-control children were not receiving additional instruction in other languages taught in schools (for the former because of their language impairment and for the latter because of their age/grade in school). This allowed us to control for the languages the children were exposed to and propose a homogeneous group, as far as possible, in relation to language exposure and use. Prior to the study, the children with SLI were assessed on a large test battery by certified speech and language therapists, including the second author. To qualify, children had to score lower than the normal range on the standardized tests in Greek in two (or more) linguistic domains. The typically language developing children serving as language-matched controls were matched with the multilingual SLI group based on scores from the standardized Greek version (Vogindroukas et al., 2009) of the Renfrew Word Finding Vocabulary Test (Renfrew, 1997).

Demographic information of the participants and results of the SLI and TLD groups on our language battery are presented in **Table 4**.

At a glance, the results from the two clinical and the two control groups can be depicted as in **Figure 5**.

The four groups were simultaneously compared on the two dependent variables (percentage correct on nouns and percentage correct on verbs), using the non-parametric Kruskal-Wallis test, which revealed significant mean differences on noun and verb accuracies [$\chi^2_{(3)} = 18.132$, $p < 0.001$ and $\chi^2_{(3)} = 27.422$, $p < 0.001$, correspondingly]. Pairwise comparisons of the multiSLI group with the other three groups were conducted with Mann-Whitney U-tests, adopting a Bonferroni adjusted level of significance (0.05/3 = 0.017). When naming accuracies for verbs and nouns of the multiSLI group were compared with the performance of the bilSLI children, the difference was not statistically significant for either word class ($z = -0.604$, $p = 0.546$ for nouns and $z = -0.698$, $p = 0.485$ for verbs). Similarly, when performance of the multiSLI group was compared to that of their multiTLD-LM peers, there was not a statistically significant difference in naming nouns ($z = -0.123$, $p = 0.902$) or verbs ($z = 0$, $p = 1$). Also, the multiSLI group scored considerably lower than the bilTLD-LM, but the difference failed to reach the adjusted level of significance ($z = -2.185$, $p = 0.029$ for nouns; $z = -2.081$, $0 = 0.037$ for verbs).
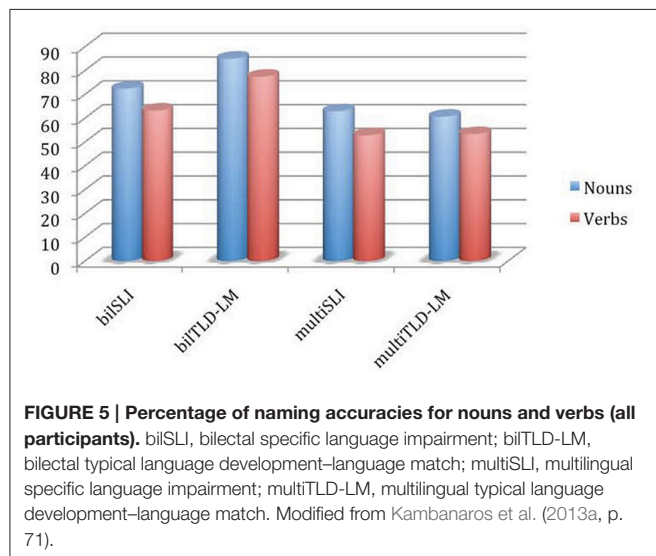
For the multilingual groups in particular, a Wilcoxon signed ranks tests was used to compare naming accuracy for nouns vs. verbs. Performance on nouns was significantly higher than for verbs for the multiSLI group ($z = -2.023$, $p = 0.043$); noun accuracy was higher than verbs but not significantly so for the multiTLD-LM group ($z = -1.070$, $p = 0.285$). Paired t-tests results concurred with the non-parametric ones. The three English-speaking multilingual children were further tested in English and all showed a better performance in their L2 (SMG) compared to their L1b (English), arguably bootstrapped by their close native L1a (CG); noun accuracy was higher than verb accuracy in both languages.

TABLE 4 | Performance on background measures (by group).

		TLD scores (SD)	MultiSLI scores (SD)	bilSLI scores (SD)
Raven's matrices		94.58 (9.64)	90 (10)	85.4 (8.89)
DVIQ—Production of Morphosynta	TLD scores are the mean results of administrating the test to a subset of 16 bilTLD children aged 4;6–9;11	19.9 (2.11)	16.8 (4.66)*	12.3 (2.09)*
DVIQ—Comprehension of Morphosyntax		26.4 (2.46)	22.6 (4.28)*	22.4 (1.84)*
DVIQ—Sentence Repetition		46.8 (1.80)	42.8 (2.77)*	40.8 (2.70)*
DVIQ—Vocabulary		22.3 (1.58)	19.04 (2.41)*	15.7 (2.20)*
DVIQ—Metalinguistic abilities		20.1 (2.45)	19.8 (1.79)	17.5 (1.29)
Word Finding Vocabulary Test (WFVT)	Norms (in Greek) for children aged 5;1–8;0 26.5–33.2	30.2 (7.9)	24.4 (4.77)	27.5 (4.96)
	10 multiTLD (mean age 5;1)	23.6 (5.87)		

*Impaired; bilSLI, bilectal SLI; multiSLI, multilingual SLI; SD, standard deviation; TLD, typical language development.

FIGURE 5 | Percentage of naming accuracies for nouns and verbs (all participants). bilSLI, bilectal specific language impairment; bilTLD-LM, bilectal typical language development–language match; multiSLI, multilingual specific language impairment; multiTLD-LM, multilingual typical language development–language match. Modified from Kambanaros et al. (2013a, p. 71).

Of the types of errors that were coded, multilingual children with and without SLI made more errors overall than typically developing bilectal children for both nouns and verbs. Omission errors for nouns also appear more frequently in both multilingual groups, where the multiSLI made more verb semantic errors and the multiTLD-LM more verb omission errors. In non-parametric group comparisons on each type of error, it was found that the groups differ significantly on noun and verb omission errors $[\chi^2_{(3)} = 16.615, p = 0.001$ and $\chi^2_{(3)} = 18.083, p < 0.001]$ as well as verb semantic errors $[\chi^2_{(3)} = 17.948, p < 0.001]$. Further pairwise comparisons revealed that the two multilingual groups made significantly more omission and verb semantic errors than the typically developing bilectal children. In essence, error type did not distinguish SLI groups (bilectal vs. multilingual).

In sum, multilingual children with SLI, like their monolingual and bilectal language-impaired peers, perform analogously to language-matched children on naming accuracy for verbs and noun on a picture-based naming task. Once more, verbs are

significantly more difficult to retrieve than nouns—a finding comparable to the monolingual and bilectal studies conducted so far in the literature (Kambanaros et al., 2013a). Taken together, these data points substantiate the claim that children with SLI, irrespective of whether they are monolingual, bilectal, or multilingual, demonstrate: (i) lexical (word-level) skills similar to younger counterparts with typical language development; (ii) no evidence of deviant or disrupted acquisition in (at least) the lexical domain; (iii) a significantly greater difficulty in retrieving verbs as opposed to nouns; (iv) consistency of omissions as the major error type for nouns across languages; and (v) divergence in the major error type for verbs across languages. This is an issue for the role of language proximity in (impaired) language development, whichever direction it is going to be implemented: Multilingual children do not show different, perhaps "additional," problems compared to bilingual ones, regardless of the additional language(s)—and not compared to the closely related bilectals either.

Our findings thus constitute the first indication from multilingual children with SLI in support of the *delayed acquisition hypothesis* for SLI (Rice, 2003). The relevance of this becomes obvious once the next step is considered in a language-impaired child's development: appropriate intervention or speech–language therapy. One major issue for speech–language therapists is how to go about treating (multilingual) children with SLI. In a related recent study (Kambanaros et al., 2015), we reported on lexical retrieval deficits using an equivalent-based measure of expressive vocabulary in the three languages of a multilingual school-aged child diagnosed with SLI. In follow-up work (Kambanaros et al., submitted), we carried out a therapy study treating cognates in one of the child's three languages (English) and observed an effect in her other two languages (Bulgarian and Greek).

DISCUSSION AND OUTLOOK

Addressing the present *Frontiers* research topic, we take "the grammar of multilingualism" to be a highly complex area of

research that by definition needs to include a lot of different measurements—by which we mean, ideally, the investigation of different measures, different sets of data, different populations, all carried out by interdisciplinary research teams. There is a need for thorough sociolinguistic work, putting the languages under investigation into their social and communicative context, for example. There is a need for thorough theoretical linguistic work, identifying the relevant structures and patterns to be investigated. There is a need for thorough psycholinguistic work, designing and carrying out the best possible experimental methodology. There is a need for cognitive psychological work, probing executive control abilities. And there is a need for clinical linguistic work, assessing and treating language impairment.

This list can be added to and enriched in many ways. The bottom line is that the notion of *comparative bilingualism* can be quite useful and instructive for future research activities, especially when carried out across different countries and languages. The narrow goal of this article was thus to draw attention to this state of affairs and elaborate the research path of comparative bilingualism (Grohmann, 2014b), with a focus on Cyprus (Grohmann and Leivada, 2012, 2013; Kambanaros et al., 2013b; Rowe and Grohmann, 2013, 2014; Karpava and Grohmann, 2014). One such intriguing path would be the role of comparative bilingualism for children with developmental language impairment, something we pointed to as well (Kambanaros et al., 2013a, 2014, 2015), even for therapy strategies (Kambanaros et al., submitted).

However, there is also a broader, larger message behind the above. We could only touch on the role of atypical and impaired language development, and only hint at further comparisons with acquired language disorders and language breakdown in age. A particular avenue of research that investigates more closely the commonalities behind these may be couched within what Benítez-Burraco and Boeckx (2014) refer to as *comparative biolinguistics*, that "inter- and intra-species variation that lies well beneath the surface variation that is the bread and butter of comparative linguistics" (Boeckx, 2013, pp. 5–6). This is a larger research enterprise, continuing the list started above. The primary aim is to obtain distinctive linguistic profiles regarding lexical and grammatical abilities, concomitant with the goal to develop cognitive profiles such as executive control across a range of genetically and non-genetically different populations who are bilectal and multilingual, with or without co-morbid linguistic and/or cognitive impairments as part of their genotype. While individual variability is clinically crucial, population-based research can advance cognitive–linguistic theory through behavioral testing that acknowledges the brain bases involved. This will offer a unique opportunity to researchers in cognitive neuroscience, psychology, speech and language therapy/pathology, psycho- and neurolinguistics, and language development to collaborate.

Our more immediate and local hope is to integrate such research backgrounds within CAT, since we believe that Cyprus is predestined to carry out such population-based research rather easily, at least from a logistical perspective: Cyprus is a small country, hosts many different cultural and linguistic backgrounds, has bilectal, bi-, and multilingual speakers, and much of what we report for the Greek-speaking Republic of Cyprus also transfers, almost mirror-like, to the Turkish-speaking occupied northern part of the island; in addition, despite its limited geographical size and population numbers, all relevant and, for clinical linguistic purposes, "interesting" disorders can be found on the island, be it genetic malfunctions, developmental impairments, or acquired disorders. In reality, however, this kind of research could, and should, be picked up anywhere in the world.

For such research, children with developmental language disorders that are language-/behavior-based or as the result of a genetic syndrome should be targeted, including the following pathological conditions which we know exist in Cyprus in research-appropriate numbers:

- Specific Language Impairment (SLI): SLI is considered a language disorder in children exhibiting difficulties acquiring grammar, phonological skills, semantic knowledge, and vocabulary, despite having a non-verbal IQ within the normal range.
- Developmental Dyslexia (DD): Children with DD experience problems learning to read, write, and spell below their chronological age, despite having a non-verbal IQ within the normal range.
- Autism Spectrum Disorders (ASD): Children with a high-functioning ASD, such as Asperger's, have problems with language and communication; they also show repetitive and/or restrictive patterns and thoughts of behavior, despite having a non-verbal IQ within the normal range.
- Down Syndrome (DS): DS is caused by three instead of the normal two copies of chromosome 21; children present with language and cognitive deficits, though differently from WS.
- William's Syndrome (WS): Individuals with WS miss ±28 genes from one copy of chromosome 7; children present with language and cognitive deficits, though differently from DS.
- Fragile X Syndrome (FXS): In FXS a particular piece of genetic code has been multiplied several times on one copy of the X chromosome; children present with language and cognitive deficits.

Each clinical (child) population provides a different perspective on language acquisition and impairment in terms of the relative strengths and weaknesses of certain processes or abilities based on the etiology and are defined as *primary language delay*, where non-linguistic cognitive skills are developing normally (here: SLI and DD), and *secondary language delay*, where language problems are secondary to other conditions (here: ASD, DS, WS, FXS). Statistical procedures can be used to compute and correlate relationships between the research measures and combinations of the background/selection markers. The results will provide new directions for investigating language impairments by considering a broad range of linguistic, cognitive, and behavioral indicators in the realm of bilectalism and multilingualism. This will also allow both associations and dissociations to emerge, and the identification of which factors co-vary with performance scores. In simple terms, it will enable us to understand the "*how*" and the "*why*" of child differences from one to another within and

across clinical conditions, and as compared to non-impaired populations.

Putting all of this together, though, there is an even more general issue. Comparing cognitive and linguistic abilities across different populations and different groups of speakers may ask for a further "specialized" area of research. The intention is to compare linguistic and cognitive abilities of monolingual, bidialectal, bilectal, bilingual, and multilingual speakers (comparative bilingualism, with more room for gradience, especially in combination such as Russian–Greek bilinguals in Cyprus) and different language-impaired populations (comparative biolinguistics, unearthing phenotypal variation), who themselves may be on different scales in the gradient spectrum of multilingualism. That is, among the future research participants, there will be vast variation and combinations of "lingual" features, ranging from mono- to multilingualism, from simultaneous to sequential acquisition, from local

to heritage language status, from typical development to impairment, from healthy to disorders of various degrees. We tentatively suggest a(nother) new term for this and are excited about what future research may bring: *comparative linguality*.

AUTHOR CONTRIBUTIONS

Both authors made substantial, direct, and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

REFERENCES

Adesope, O. O., Lavin, T., Thompson, T., and Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Rev. Educ. Res.* 80, 207–245. doi: 10.3102/0034654310368803

Agathocleous, M. (2012). *The Role of the Social Environment in the Production of Clitics in Cypriot Greek within the COST Action A33 Testing Tool.* MA thesis, University of Cyprus, Nicosia.

Agathocleous, M., Charalambous, A., Papadopoulou, E., and Grohmann, K. K. (2014). "The role of the social environment on linguistic development: a view from Cypriot Greek clitic placement," in *Developments in the Acquisition of Clitics*, eds K. K. Grohmann and T. Neokleous (Newcastle: Cambridge Scholars Publishing), 41–86.

Agouraki, Y. (1997). "On the enclisis/proclisis alternation," in *Proceedings of the Second International Conference on Greek Linguistics*, Vol. II, eds G. Drachman, A. Malikouti-Drachman, J. Fykias, and C. Klidi (Salzburg: University of Salzburg), 393–404.

Aikhenvald, A. Y. (2007). "Grammars in contact: a cross-linguistic perspective," in *Grammars in Contact: A Cross-Linguistic Typology*, eds A. Y. Aikhenvald and R. M. W. Dixon (Oxford: Oxford University Press), 1–66.

Andreou, M. (2015). *The Effects of Bilingualism on Verbal and Non-Verbal Cognition: The Micro- and Macro-Structure of Narratives in the Weak and the Dominant Language of a Bilingual Child.* Ph.D. dissertation, Aristotle University of Thessaloniki.

Antoniou, K. (2014). *The Effects of Childhood Bilectalism and Bilingualism on Executive Control and Implicature Understanding.* Ph.D. dissertation, University of Cambridge, Cambridge.

Antoniou, K., Grohmann, K. K., Kambanaros, M., and Katsos, N. (2016). The effect of childhood bilectalism and multilingualism on executive control. *Cognition* 149, 18–30. doi: 10.1016/j.cognition.2015.12.002

Antoniou, K., Kambanaros, M., Grohmann, K. K., and Katsos, N. (2014). "Is bilectalism similar to bilingualism? An investigation into children's vocabulary and executive control skills," in *BUCLD* 38: *Proceedings of the 38th Annual Boston University Conference on Language Development*, Vol. 1, eds W. Orman and M. J. Valleau (Somerville, MA: Cascadilla Press), 12–24.

Arvaniti, A. (2010). Linguistic practices in Cyprus and the emergence of Cypriot Standard Greek. *Mediterr. Lang. Rev.* 17, 15–45.

Baker, C. (1996). *Foundations of Bilingual Education and Bilingualism*, 2nd Edn. Clevedon: Multilingual Matters.

Barac, R., and Bialystok, E. (2012). Bilingual effects on cognitive and linguistic development: role of language, cultural background, and education. *Child Dev.* 83, 423–422. doi: 10.1111/j.1467-8624.2011.01707.x

Barac, R., Bialystok, E., Castro, D. C., and Sanchez, M. (2014). The cognitive development of young dual language learners: a critical review. *Early Child. Res. Q.* 29, 699–714. doi: 10.1016/j.ecresq.2014.02.003

Baum, S., and Titone, D. (2014). Moving towards a neuroplasticity view of bilingualism, executive control, and aging. *Appl. Psycholinguist.* 35, 857–894. doi: 10.1017/S0142716414000174

Benítez-Burraco, A., and Boeckx, C. (2014). Universal Grammar and biological variation: an EvoDevo agenda for comparative biolinguistics. *Biol. Theory* 9, 122–134. doi: 10.1007/s13752-014-0164-0

Bernbaum, G. (1979). *Comparative Bilingualism*. Cambridge, MA: National Assessment and Dissemination Center for Bilingual/Bicultural Education.

Bialystok, E. (1997). The structure of age: in search of barriers to second language acquisition. *Second Lang. Res.* 13, 116–137. doi: 10.1191/026765897677670241

Bialystok, E. (2001). *Bilingualism in Development: Language, Literacy, and Cognition*. Cambridge: Cambridge University Press.

Bialystok, E. (2009). Bilingualism: the good, the bad, and the indifferent. *Bilingualism Lang. Cogn.* 12, 3–11. doi: 10.1017/S1366728908003477

Bialystok, E., Craik, F., and Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends Cogn. Sci.* 16, 240–250. doi: 10.1016/j.tics.2012.03.001

Bialystok, E., and Viswanathan, M. (2009). Components of executive control with advantages for bilingual children in two cultures. *Cognition* 112, 494–500. doi: 10.1016/j.cognition.2009.06.014

Birdsong, D., and Molis, M. (2001). On the evidence of maturational constraints in second language acquisition. *J. Mem. Lang.* 44, 235–249. doi: 10.1006/jmla.2000.2750

Blom, E., Küntay, A. C., Messer, M., Verhagen, J., and Leseman, P. (2014). The benefits of being bilingual: working memory in bilingual Turkish–Dutch children. *J. Exp. Child Psychol.* 128, 105–119. doi: 10.1016/j.jecp.2014.06.007

Boeckx, C. (2013). Biolinguistics: forays into human cognitive biology. *J. Anthropol. Sci.* 91, 1–28.

Calvo, A., and Bialystok, E. (2014). Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. *Cognition* 130, 278–288. doi: 10.1016/j.cognition.2013.11.015

Cazden, C. B. (1976). "Play with language and metalinguistic awareness: one dimension of language experience," in *Play: Its Role in Development and Evolution*, eds J. S. Bruner, A. Jolly, and K. Sylva (New York, NY: Basic Books), 603–408.

Ceñoz, J., and Gorter, D. (2011). Focus on multilingualism: a study of trilingual writing. *Modern Lang. J.* 95, 356–369. doi: 10.1111/j.1540-4781.2011.01206.x

Charalambous, A. (2012). *The Development of Clitic Placement in Cypriot Greek within the COST Action IS0804 Testing Tool: Factors and Trends.* MA thesis, University of Cyprus, Nicosia.

Costa, A., Hernández, M., Costa-Faidella, J., and Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: now you see it, now you don't. *Cognition* 113, 135–149. doi: 10.1016/j.cognition.2009.08.001

Costa, A., and Sebastián-Gallés, N. (2014). How does the bilingual experience sculpt the brain? *Nat. Rev. Neurosci.* 15, 336–345. doi: 10.1038/nrn3709

Costa, J., and Lobo, M. (2007). "Clitic omission, null object or both in the acquisition of European Portuguese?" in *Romance Languages and Linguistic Theory 2005*, eds S. Baauw, F. Drijkoningen, and M. Pinto (Amsterdam: John Benjamins), 59–72.

COST Action A33 (2006–2010). *Cross-Linguistically Robust Stages of Children's Linguistic Performance with Applications to the Diagnosis of Specific Language Impairment*. Brussels: COST Office. Available online at: http://www.zas.gwz-berlin.de/cost.html?&L=0

COST Action IS0804 (2009–2013). *Language Impairment in a Multilingual Society: Linguistic Patterns and the Road to Assessment*. Brussels: COST Office. Available online at: http://www.bi-sli.org

Currie, C. E., Elton, R. A., Todd, J., and Platt, S. (1997). Indicators of socioeconomic status for adolescents: the WHO Health Behaviour in School-Aged Children Survey. *Health Educ. Res.* 12, 385–397. doi: 10.1093/her/12.3.385

de Abreu, P. M. E. (2011). Working memory in multilingual children: is there a bilingual effect? *Memory* 19, 529–537. doi: 10.1080/09658211.2011.590504

de Abreu, P. M. E., Cruz-Santos, A., Tourinho, C. J., Martin, R., and Bialystok, E. (2012). Bilingualism enriches the poor enhanced cognitive control in low-income minority children. *Psychol. Sci.* 23, 1364–1371. doi: 10.1177/0956797612443836

Du Bois, J. A. (1985). "Competing motivations," in *Iconicity in Syntax*, ed J. Haiman (Amsterdam: John Benjamins), 343–366.

Falk, Y., and Bardell, C. (2010). The study of the role of the background languages in third language acquisition: the state of the art. *Int. Rev. Appl. Linguist. Lang. Teach.* 48, 185–219. doi: 10.1515/iral.2010.009

Ferguson, C. A. (1959 [2003]). "Diglossia. *Word* 15, 325–340," in *Sociolinguistics: The Essential Readings*, eds C. Bratt Paulston and G. R. Tucker (Oxford: Blackwell), 345–358, 2003.

Foy, J. G., and Mann, V. A. (2014). Bilingual children show advantages in nonverbal auditory executive function task. *Int. J. Bilingualism* 18, 717–729. doi: 10.1177/1367006912472263

Friedmann, N., Biran, M., and Dotan, D. (2013). "Lexical retrieval and its breakdown in aphasia and developmental language impairment," in *The Cambridge Handbook of Biolinguistics*, eds C. Boeckx and K. K. Grohmann (Cambridge: Cambridge University Press), 350–374.

Genesee, F., Paradis, J., and Crago, M. B. (2011). *Dual Language Development and Disorders: A Handbook on Bilingualism and Second Language Learning, 2nd Edn*. Baltimore, MD: Paul Brookes.

German, D. J., and Newman, R. S. (2004). The impact of lexical factors on children's word-finding errors. *J. Speech Lang. Hear. Res.* 47, 624–636. doi: 10.1044/1092-4388(2004/048)

Grohmann, K. K. (2009). "Focus on clefts: a perspective from Cypriot Greek," in *Selected Papers from the 18th International Symposium on Theoretical and Applied Linguistics—Thessaloniki, 4–6 May 2007*, ed A. Tsangalides (Thessaloniki: Monochromia), 157–165.

Grohmann, K. K. (2011). "Some directions for the systematic investigation of the acquisition of Cypriot Greek: a new perspective on production abilities from object clitic placement," in *The Development of Grammar: Language Acquisition and Diachronic Change*, eds E. Rinke and T. Kupisch (Amsterdam: John Benjamins), 179–203.

Grohmann, K. K. (2014a). "CAT research on object clitic placement: where we are now," in *Developments in the Acquisition of Clitics*, eds K. K. Grohmann and T. Neokleous (Newcastle: Cambridge Scholars Publishing), 1–40.

Grohmann, K. K. (2014b). Towards comparative bilingualism. *Linguist. Approaches Bilingualism* 4, 336–341. doi: 10.1075/lab.4.3.06gro

Grohmann, K. K., and Leivada, E. (2011). "Towards disentangling bi-x: metalinguistic awareness and competing motivations," paper presented at the *2011 UIC Bilingualism Forum*, University of Illinois at Chicago (14–15 April 2011).

Grohmann, K. K., and Leivada, E. (2012). "Interface ingredients of dialect design: Bi-*x*, socio-syntax of development, and the grammar of Cypriot Greek," in *Towards a Biolinguistic Understanding of Grammar: Essays on Interfaces*, ed A. M. Di Sciullo (Amsterdam: John Benjamins), 239–262.

Grohmann, K. K., and Leivada, E. (2013). "Lightverbhood in child language: evidence from Cypriot Greek," in *Dummy Auxiliaries in First and Second Language Acquisition*, eds E. Blom, J. Verhagen, and I. van de Craats (Berlin: De Gruyter Mouton), 141–170.

Grohmann, K. K., Panagiotidis, P., and Tsiplakou, S. (2006). "Properties of wh-question formation in Cypriot Greek," in *Proceedings of the 2nd International Conference on Modern Greek Dialects and Linguistic Theory (Mytilene, Greece: 30 September–3 October 2004)*, eds M. Janse, B. D. Joseph, and A. Ralli (Patras: University of Patras), 83–98.

Grohmann, K. K., and Papadopoulou, E. (2011). Question(able) issues in Cypriot Greek. *Linguist. Anal.* 37, 8–38.

Grohmann, K. K., Theodorou, E., Pavlou, N., Leivada, E., Papadopoulou, E., and Martínez-Ferreiro, S. (2012). "The development of object clitic placement in Cypriot Greek and the Romance connection," in *Selected Proceedings of the Romance Turn IV Workshop on the Acquisition of Romance Languages*, eds S. Ferré, P. Prévost, L. Tuller, and R. Zebib (Newcastle-upon-Tyne: Cambridge Scholars Publishing), 128–152.

Hadjioannou, X., Tsiplakou, S., and Kappler, M. (2011). Language policy and language planning in Cyprus. *Curr. Issues Lang. Plann.* 12, 503–569. doi: 10.1080/14664208.2011.629113

Ivanova, I., and Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychol.* 127, 277–288. doi: 10.1016/j.actpsy.2007.06.003

Jarvis, S., and Pavlenko, A. (2008). *Crosslinguistic Influence in Language and Cognition*. London: Routledge.

Kambanaros, M. (2003). *Verb and Noun Processing in Late Bilingual Individuals with Anomic Aphasia*. Ph.D. dissertation, Flinders University of South Australia, Adelaide.

Kambanaros, M. (2013). Does verb type affect action naming in specific language impairment (SLI)? Evidence from instrumentality and name relation. *J. Neurolinguist.* 26, 160–177. doi: 10.1016/j.jneuroling.2012.07.003

Kambanaros, M., and Grohmann, K. K. (2010). "Patterns of object and action naming in Cypriot Greek children with SLI and WFDs," in *BUCLD 34: Proceedings of the 34th Annual Boston University Conference on Language Development—Supplement*, eds K. Franich, L. Keil, K. Iserman, and J. Chandlee. Available online at: http://www.bu.edu/bucld/proceedings/supplement/vol34)

Kambanaros, M., and Grohmann, K. K. (2011). "Patterns of naming objects and actions in Cypriot Greek children with SLI and WFDs," in *Selected Papers from the 19th International Symposium on Theoretical and Applied Linguistics—Thessaloniki, 3–5 April 2009*, ed A. Tsangalides (Thessaloniki: Monochromia), 233–242.

Kambanaros, M., and Grohmann, K. K. (2015). More GAPs in children with SLI? Evidence from Greek for not fully lexical verbs in language development. *Appl. Psycholinguist.* 36, 1029–1057. doi: 10.1017/S0142716414000034

Kambanaros, M., Grohmann, K. K., and Michaelides, M. (2013b). Lexical retrieval for nouns and verbs in typically developing bilectal children. *First Lang.* 33, 182–199. doi: 10.1177/0142723713479435

Kambanaros, M., Grohmann, K. K., Michaelides, M., and Theodorou, E. (2013a). Comparing multilingual children with SLI to their bilectal peers: evidence from object and action picture naming. *Int. J. Multilingualism* 10, 60–81. doi: 10.1080/14790718.2012.705846

Kambanaros, M., Grohmann, K. K., Michaelides, M., and Theodorou, E. (2014). On the nature of verb–noun dissociations in bilectal SLI: a psycholinguistic perspective from Greek. *Bilingualism Lang. Cogn.* 17, 169–188. doi: 10.1017/S1366728913000035

Kambanaros, M., Grohmann, K. K., and Theodorou, E. (2010). "Action and object naming in mono- and bilingual children with specific language impairment," in *Proceedings of ISCA Tutorial and Research Workshop on Experimental Linguistics 2010—25–27 August 2010, Athens, Greece*, ed A. Botinis (Athens: ISCA and University of Athens), 73–76.

Kambanaros, M., Michaelides, M., and Grohmann, K. K. (2015). Measuring word retrieval deficits in a multilingual child with SLI: is there a better language? *J. Neurolinguist.* 34, 112–130. doi: 10.1016/j.jneuroling.2014.09.006

Karpava, S. (2014). *Vulnerable Domains for Cross-Linguistic Influence in L2 Acquisition of Greek*. Ph.D. dissertation, University of Cyprus, Nicosia.

Karpava, S., and Grohmann, K. K. (2014). "Bilingual acquisition of Cypriot Greek object clitic placement," in *Developments in the Acquisition of Clitics*, eds K. K. Grohmann and T. Neokleous (Newcastle: Cambridge Scholars Publishing), 87–137.

Kovács, Á. M., and Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants. *Proc. Nat. Acad. Sci.* 106, 6556–6560. doi: 10.1073/pnas.0811323106

Kroch, A. (1994). "Morphosyntactic variation," in *Papers from the 30th Regional Meeting of the Chicago Linguistics Society: parasession on Variation and*

Linguistic Theory, eds K. Beals, J. Denton, R. Knippen, L. Melnar, H. Suzuki, and E. Zeinfeld (Chicago: Chicago Linguistics Society), 180–201.

Kroch, A., and Taylor, A. (2000). "Verb–object order in Middle English," in *Diachronic Syntax: Models and Mechanisms*, eds S. Pintzuk, G. Tsoulas, and A. Warner (Oxford: Oxford University Press), 132–163.

Kroll, J. F., and Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *J. Cogn. Psychol.* 25, 497–514. doi: 10.1080/20445911.2013.799170

Legate, J., and Yang, C. (2007). Morphosyntactic learning and the development of tense. *Lang. Acquis.* 14, 315–344. doi: 10.1080/10489220701471081

Leivada, E., and Grohmann, K. K. (2016). "Language acquisition in bilectal environments: Competing motivations, metalinguistic awareness, and the socio-syntax of development hypothesis," in *Variation in Language Acquisition*, eds M. Katerbow and G. de Vogelaer (Amsterdam: John Benjamins).

Leivada, E., Mavroudi, P., and Epistithiou, A. (2010). "Metalanguage or bidialectism? Acquisition of clitic placement by Hellenic Greeks, Greek Cypriots and binationals in the diglossic context of Cyprus," in *Proceedings of ISCA Tutorial and Research Workshop on Experimental Linguistics 2010*, ed A. Botinis (Athens: ISCA and the University of Athens), 97–100.

Lewis, M. P., Simons, G. F., and Fennig, C. D. (eds.). (2015). *Ethnologue: Languages of the World, 18th Edn.* Dallas, TX: SIL International. Available online at: http://www.ethnologue.com.

Lightfoot, D. (1999). *The Development of Language: Acquisition, Change, and Evolution*. Oxford: Blackwell.

Llama, R., Cardoso, W., and Collins, L. (2009). The influence of language distance and language status on the acquisition of L3 phonology. *Int. J. Multilingualism* 7, 39–57. doi: 10.1080/14790710902972255

Marinis, T. (2000). The acquisition of clitic objects in Modern Greek: single clitics, clitic doubling, clitic left dislocation. *ZAS Working Papers* 15, 259–281.

Marx, N., and Mehlhorn, G. (2010). Pushing the positive: encouraging phonological transfer from L2 to L3. *Int. J. Multilingualism* 7, 4–18. doi: 10.1080/14790710902972271

Meisel, J. M. (2007). The weaker language in early childhood bilingualism: acquiring a first language as a second language. *Appl. Psycholinguist.* 28, 495–514. doi: 10.1017/S0142716407070270

Meisel, J. M. (2009). Second language acquisition in early childhood. *Zeitschrift für Sprachwissenschaft* 28, 5–34. doi: 10.1515/zfsw.2009.002

Messer, D., and Dockrell, J. (2011). Lexical access and literacy in children with word-finding difficulties. *Int J. Lang. Commun. Disord.* 46, 473–480. doi: 10.1111/j.1460-6984.2011.00009.x

Miyake, A., and Friedman, N. P. (2012). The nature and organization of individual differences in executive functions four general conclusions. *Curr. Dir. Psychol. Sci.* 21, 8–14. doi: 10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex 'frontal lobe' tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Montrul, S. (2008). *Incomplete Acquisition in Bilingualism: Re-examining the Age Factor*. Amsterdam: John Benjamins.

Morales, J., Calvo, A., and Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *J. Exp. Child Psychol.* 114, 187–202. doi: 10.1016/j.jecp.2012.09.002

Muller, J. (2002). On the theoretical implications of Cypriot Greek initial geminates. *J. Greek Linguist.* 3, 115–137. doi: 10.1075/jgl.3.06mul

Newton, B. E. (1972). *Cypriot Greek: Its Phonology and Inflections*. The Hague: Mouton.

Oller, D., and Eilers, R. (2002). *Language Literacy in Bilingual Children*. Clevedon: Multilingual Matters.

Paap, K. R., and Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cogn. Psychol.* 66, 232–258. doi: 10.1016/j.cogpsych.2012.12.002

Paap, K. R., and Sawi, O. (2014). Bilingual advantages in executive functioning: problems in convergent validity, discriminant validity, and the identification of the theoretical constructs. *Front. Psychol. Lang. Sci.* 5:962. doi: 10.3389/fpsyg.2014.00962

Papadopoulou, E., Leivada, E., and Pavlou, N. (2014). Acceptability judgments in bilectal populations: competition, gradience, and socio-syntax. *Linguist. Variation* 14, 109–128. doi: 10.1075/lv.14.1.05pap

Paradis, J. (2011). Individual differences in child English second language acquisition: comparing child-internal and child-external factors. *Linguist. Approaches Bilingualism* 1, 213–237. doi: 10.1075/lab.1.3.01par

Paradis, J., Emmerzael, K., and Duncan, T. S. (2010). Assessment of English language learners: using parent report on first language development. *J. Commun. Disord.* 43, 474–497. doi: 10.1016/j.jcomdis.2010.01.002

Paradis, J., and Genesee, F. (1996). Syntactic acquisition in bilingual children: autonomous or independent? *Stud. Second Lang. Acquis.* 18, 1–25. doi: 10.1017/S0272263100014662

Parodi, T., and Tsimpli, I. M. (2005). 'Real' and apparent optionality in second language grammars: finiteness and pronouns in null operator structures. *Second Lang. Res.* 21, 250–285. doi: 10.1191/0267658305sr248oa

Petinou, K., and Terzi, A. (2002). Clitic misplacement among normally developing children and children with specific language impairment and the status of Infl heads. *Lang. Acquis.* 10, 1–28. doi: 10.1207/S15327817LA1001_1

Poarch, G. J., and van Hell, J. G. (2012). Executive functions and inhibitory control in multilingual children: evidence from second-language learners, bilinguals, and trilinguals. *J. Exp. Child Psychol.* 113, 535–551. doi: 10.1016/j.jecp.2012.06.013

Prior, A., and Gollan, T. H. (2011). Good language-switchers are good task-switchers: evidence from Spanish–English and Mandarin–English bilinguals. *J. Int. Neuropsychol. Soc.* 17, 682–691. doi: 10.1017/S1355617711000580

Renfrew, C. (1997). *The Renfrew Language Scales —Bus Story Test: A Test of Narrative Speech/Word Finding Vocabulary Test/Action Picture Test*, 4th Edn. Milton Keynes: Speechmark.

Rice, M. L. (2003). "A unified model of specific and general language delay: grammatical tense as a clinical marker of unexpected variation," in *Language Competence across Populations: Toward a Definition of Specific Language Impairment*, eds Y. Levy and J. Schaeffer (Mahwah, NJ: Lawrence Erlbaum Associates), 63–95.

Ringbom, H. (2006). *Cross-Linguistic Similarity in Foreign Language Learning*. Clevedon: Multilingual Matters.

Ringbom, H., and Jarvis, S. (2009). "The importance of cross-linguistic similarity in foreign language learning," in *The Handbook of Language Teaching*, eds M. H. Long and C. J. Doughty (Oxford: Wiley-Blackwell), 106–118.

Rowe, C., and Grohmann, K. K. (2013). Discrete bilectalism: towards co-overt prestige and diglossic shift in Cyprus. *Int. J. Soc. Lang.* 224, 119–142. doi: 10.1515/ijsl-2013-0058

Rowe, C., and Grohmann, K. K. (2014). Canaries in a coal mine: native speakerhood and other factors as predictors of moribundity, death, and diglossic shift in Cypriot Greek. *Mediterr. Lang. Rev.* 21, 121–142.

Runnqvist, E., FitzPatrick, I., Strijkers, K., and Costa, A. (2012). "An appraisal of the bilingual language production system: quantitatively or qualitatively different from monolinguals?" in *The Handbook of Bilingualism and Multilingualism, 2nd Edn.*, eds T. K. Bhatia and W. C. Ritchie (Malden, MA: Wiley-Blackwell), 244–265.

Siegal, M., Iozzi, L., and Surian, L. (2009). Bilingualism and conversational understanding in young children. *Cognition* 110, 115–122. doi: 10.1016/j.cognition.2008.11.002

Siegal, M., Matsuo, A., Pond, C., and Otsu, Y. (2007). "Bilingualism and cognitive development: evidence from scalar implicatures," in *Proceedings of the Eighth Tokyo Conference on Psycholinguistics*, ed Y. Otsu (Tokyo: Hituzi Syobo), 265–280.

Siegal, M., Surian, L., Matsuo, A., Geraci, A., Iozzi, L., Okumura, Y., et al. (2010). Bilingualism accentuates children's conversational understanding. *PLoS ONE* 5:e9004. doi: 10.1371/journal.pone.0009004

Siegel, J. (2010). *Second Dialect Acquisition*. Cambridge: Cambridge University Press.

Sophocleous, A. (2011). Two languages in the classroom: the inconsistency between national and local objectives in formal education in Cyprus. *J. Lang. Identity and Educ.* 10, 266–281. doi: 10.1080/15348458.2011.598129

Sophocleous, A., and Wilks, C. (2010). Standard Modern Greek and Greek-Cypriot dialect in kindergarten classroom interaction: teachers' and learners' language attitudes and language use. *Lang. Cult. Curriculum* 23, 51–69. doi: 10.1080/07908311003632519

Terkourafi, M. (2007). Perceptions of difference in the Greek sphere: the case of Cyprus. *J. Greek Linguist.* 8, 60–96. doi: 10.1075/jgl.8.06ter

Theodorou, E. (2013). *Specific Language Impairment in Cypriot Greek: Diagnostic and Experimental Investigations*. Ph.D. dissertation, University of Cyprus, Nicosia.

Theodorou, E., and Grohmann, K. K. (2013). The acquisition of relative clauses in Cypriot Greek: production and comprehension. *Diacrítica* 26, 271–300.

Theodorou, E., and Grohmann, K. K. (2015). Object clitics in Cypriot Greek children with SLI. *Lingua* 161, 144–158. doi: 10.1016/j.lingua.2014.11.011

Thomason, S. G. (2001). *Language Contact: An Introduction*. Edinburgh: Edinburgh University Press.

Thomason, S. G., and Kaufman, T. (1988). *Language Contact, Creolization, and Genetic Linguistics*. Berkeley, CA: University of California Press.

Tomblin, J. B. (2008). "Adolescent outcomes of SLI," in *Developmental Language Impairment*, eds C. Norbury, J. B. Tomblin, and D. V. M. Bishop (London: Psychology Press), 93–114.

Tsimpli, I. M. (2003). "Clitics and articles in L2 Greek," in *Proceedings of the 6th Generative Approaches to Second Language Acquisition Conference (GASLA 2002)*, eds J. M. Liceras, H. Zobl, and H. Goodluck (Somerville, MA: Cascadilla Proceedings Project), 331–339.

Tsimpli, I. M. (2014). Early, late or very late: timing acquisition and bilingualism. *Linguist. Approaches Bilingualism* 4, 393–402. doi: 10.1075/lab.4.3.17tsi

Tsimpli, I. M., and Mastropavlou, M. (2007). "Feature interpretability in L2 acquisition and SLI: greek clitics and determiners," in *The Role of Formal Features in Second Language Acquisition*, eds J. M. Liceras, H. Zobl, and H. Goodluck (London: Routledge), 142–183.

Tsiplakou, S. (2006). The emperor's old clothes: linguistic diversity and the redefinition of literacy. *Int. J. Hum.* 2, 2345–2352.

Tsiplakou, S. (2009). "Language alternation, competing grammars and interlanguage: another facet of the Cyprus problem," in *Proceedings of the 8th International Conference on Greek Linguistics*, eds G. Giannakis, M. Baltazani, G. Xydopoulos, and T. Tsangalidis (Ioannina: University of Ioannina), 1195–1209.

Tsiplakou, S. (2014). How mixed is a 'mixed' system? The case of the Cypriot Greek koiné. *Linguist. Variation* 14, 161–178. doi: 10.1075/lv.14.1.07tsi

Tuller, L. (2015). "Clinical use of parental questionnaires in multilingual contexts," in *Assessing Multilingual Children: Disentangling Bilingualism from Language Impairment*, eds S. Armon-Lotem, J. de Jong, and N. Meir (Bristol: Multilingual Matters), 301–330.

Tuller, L., Delage, H., Monjauze, C., Piller, A. C., and Barthez, M. A. (2011). Clitic pronoun production as a measure of atypical language development in French. *Lingua* 121, 423–441. doi: 10.1016/j.lingua.2010.10.008

Unsworth, S. (2013). Assessing age of onset effects in (early) child L2 acquisition. *Lang. Acquis.* 20, 74–92. doi: 10.1080/10489223.2013.766739

Unsworth, S., Argyri, F., Cornips, L., Hulk, A., Sorace, A., and Tsimpli, I. M. (2014). On the role of age of onset and input in early child bilingualism in Greek and Dutch. *Appl. Psycholinguist.* 35, 765–805. doi: 10.1017/S0142716412000574

Vangsnes, Ø. A., Söderlund, G. B. W., and Blekesaune, M. (2015). The effect of bidialectal literacy on school achievement. *Int. J. Biling. Educ. Biling.* doi: 10.1080/13670050.2015.1051507

Varlokosta, S., Belletti, A., Costa, J., Friedmann, N., Gavarró, A., Grohmann, K. K., et al. (2015). A crosslinguistic study of the acquisition of pronoun and clitic production. *Lang. Acquis.* doi: 10.1080/10489223.2015.1028628

Varlokosta, S., Konstantzou, K., and Nerantzini, M. (2014). "On the production of direct object clitics in Greek typical development and specific language impairment: the effect of task selection," in *Developments in the Acquisition of Clitics*, eds K. K. Grohmann and T. Neokleous (Newcastle: Cambridge Scholars Publishing), 188–211.

Vogindroukas, I., Protopapas, A., and Sideris, G. (2009). *Word-Finding Test [in Greek]*. Chania: Glafki.

Yang, C. (2000). Internal and external forces in language change. *Lang. Var. Change* 12, 231–250. doi: 10.1017/S0954394500123014

Yiakoumetti, A. (2007). Choice of classroom language in bidialectal communities: to include or to exclude the dialect? *Camb. J. Edu.* 37, 51–66. doi: 10.1080/03057640601179046

Processing Coordinate Subject-Verb Agreement in L1 and L2 Greek

Maria Kaltsa[1], Ianthi M. Tsimpli[2]*, Theodoros Marinis[3] and Melita Stavrou[4]

[1] Language Development Lab, Department of Theoretical and Applied Linguistics, Aristotle University of Thessaloniki, Thessaloniki, Greece, [2] Department of Theoretical and Applied Linguistics, University of Cambridge, Cambridge, UK, [3] Department of Clinical Language Sciences, School of Psychology and Clinical Language Sciences, University of Reading, Reading, UK, [4] Department of Linguistics, School of Philology, Aristotle University of Thessaloniki, Thessaloniki, Greece

The present study examines the processing of subject-verb (SV) number agreement with coordinate subjects in pre-verbal and post-verbal positions in Greek. Greek is a language with morphological number marked on nominal and verbal elements. Coordinate SV agreement, however, is special in Greek as it is sensitive to the coordinate subject's position: when pre-verbal, the verb is marked for plural while when post-verbal the verb can be in the singular. We conducted two experiments, an acceptability judgment task with adult monolinguals as a pre-study (Experiment 1) and a self-paced reading task as the main study (Experiment 2) in order to obtain acceptance as well as processing data. Forty adult monolingual speakers of Greek participated in Experiment 1 and a hundred and forty one in Experiment 2. Seventy one children participated in Experiment 2: 30 Albanian-Greek sequential bilingual children and 41 Greek monolingual children aged 10–12 years. The adult data in Experiment 1 establish the difference in acceptability between singular VPs in SV and VS constructions reaffirming our hypothesis. Meanwhile, the adult data in Experiment 2 show that plural verbs accelerate processing regardless of subject position. The child online data show that sequential bilingual children have longer reading times (RTs) compared to the age-matched monolingual control group. However, both child groups follow a similar processing pattern in both pre-verbal and post-verbal constructions showing longer RTs immediately after a singular verb when the subject was pre-verbal indicating a grammaticality effect. In the post-verbal coordinate subject sentences, both child groups showed longer RTs on the first subject following the plural verb due to the temporary number mismatch between the verb and the first subject. This effect was resolved in monolingual children but was still present at the end of the sentence for bilingual children indicating difficulties to reanalyze and integrate information. Taken together, these findings demonstrate that (a) 10–12 year-old sequential bilingual children are sensitive to number agreement in SV coordinate constructions parsing sentences in the same way as monolingual children even though their vocabulary abilities are lower than that of age-matched monolingual peers and (b) bilinguals are slower in processing overall.

Keywords: number agreement, coordinate subjects, child bilingualism, Greek sentence processing, adult processing

INTRODUCTION

The present study examines the processing of Subject-Verb (SV) number agreement in pre-verbal and post-verbal coordinate subject constructions in Greek. Examples (1) and (2) illustrate post-verbal and pre-verbal coordinate subject constructions in Greek. Greek has morphological number agreement (singular and plural) marking between the subject and the verb. However, coordinate subjects are a special case because number agreement is sensitive to the position of the subject. In particular, post-verbal coordinate subjects trigger plural agreement but optionally allow for singular verbs as well, as illustrated in example (1) below. In contrast, pre-verbal coordinate subjects require plural agreement while singular number agreement on the verb gives rise to ungrammaticality (Holton et al., 1997; Spyropoulos, 2007; Kazana, 2011), as shown in example (2) below.

1. Postverbal coordinate subject
    Irthan/Irthe   o Yanis ke i Maria.
    arrived.3p/arrived.3s the Yanis and the Maria
    'Yanis and Maria arrived.'

2. Preverbal coordinate subject
    O Yanis ke i Maria irthan/*irthe.
    the Yanis and the Maria arrived.3p/arrived.3s
    'Yanis and Maria arrived.'

Agreement has been argued to be either a syntactic (Chomsky, 2001; Bošković, 2009) or an entirely post-syntactic process (Bobaljik, 2008) with Closest Conjunct Agreement (CCA) accounts identifying linear proximity as a key post-syntactic component of grammar (Benmamoun, 1996; Benmamoun et al., 2009; for a detailed analysis on locating agreement see Bhatt and Walkow, 2013). Within syntactic accounts, coordinate subject agreement has been argued to be resolved with either full or partial agreement accounts. In full agreement accounts, agreement takes place with the Coordination Phrase as a whole, while feature mismatch is resolved according to resolution rules (Corbett, 1991). In partial agreement accounts (Aoun et al., 1994), agreement takes place with the closest available conjunct; in post-verbal contexts either with the first or highest conjunct (First Conjunct Agreement, FCA) and in pre-verbal contexts with the last one (Last Conjunct Agreement, LCA). In partial agreement accounts linear order between the coordinated DPs is indirectly addressed within the syntactic component. The phenomenon of partial agreement with coordinate subjects has been attested in many unrelated languages such as Arabic (Aoun et al., 1994), Slovenian (Marušič et al., 2007), Hindi (Benmamoun, 2000), and Serbo-Croatian (Bošković, 2009, 2010). This (mis)match in number agreement patterns may be addressed in two ways; either through VP coordination with verb raising, as in (3), or through DP coordination, as in (4).

3. [INFLP ν-V [νP DP$_1$ t$_{ν-V}$] and [νP DP$_2$ e$_{ν-V}$]]
    arrived   [Yanis arrived] and [Maria arrived]

4. [DP [DP$_1$]and [DP$_2$]]

According to Spyropoulos (2007) and Aoun et al. (1994; see also Johannessen, 1998; Harbert and Bahloul, 2002 for similar analyses), number mismatch cases can be accounted for by assuming VP coordination with each conjunct being the subject of its own clause, thus triggering singular agreement there. Verb-raising to the inflection head with deletion of the two lower verb copies results in a surface order whereby the singular verb is followed by two conjoined singular DPs. All other cases, that is, pre-verbal coordinate subjects and post-verbal constructions with plural number agreement, can be accounted for by assuming DP coordination, as in (4). This suggestion is in line with Munn's (1999) phrasal analysis shown to satisfy the requirements for syntactic and semantic plurality when accounting for such agreement phenomena. Notice that the analysis which assumes VP-coordination and verb-raising is syntactically more complex than DP coordination. Specifically, the fact that (3) involves a dependency involving three copies of the verb indicates higher complexity than the structure in (4) where no movement or dependency is formed. In this respect, (4) corresponds more closely to the structure of subject-verb agreement with single, non-coordinate subjects. In addition, plural number agreement with DP coordination (i.e., (4)) is semantically unmarked since the coordinate subject is semantically plural. Finally, plural agreement generalizes over pre-verbal and post-verbal coordinate subjects, whereas the structure in (3) is an option associated with post-verbal coordinate subjects only. This suggests that plural (full) agreement should be more frequent than singular and as such it should be easier to process.

It should also be noted that coordinating a singular and a plural DP subject reduces the acceptability of the singular number option on the verb, as shown by example (5) below. Furthermore, the grammaticality of the coordinate subject with a singular and a plural number DP subject deteriorates further when the plural member precedes the singular one. Compare the examples (5) and (6) below, with one plural and one singular subject DP coordinated:

5. V + S-**sing** + S-**plural**
    ?Irthe   o Yanis ke ta pedhia.
    arrived.3s the Yanis and the kids
    'Yanis and the kids arrived.'

6. V + S-**plural** + S-**sing**
    *Irthe   ta pedhia ke o Yanis.
    arrived.3s the kids   and the Yanis
    'The kids and Yanis arrived.'

The reduced acceptability of (5) may be reflected in processing patterns and response times too, although such structures have not been investigated in the processing literature yet. The additional effects of the ordering between the singular and the plural subject in (6) increase the variables that number agreement might be sensitive to in coordinate subject processing. Thus, our experimental study examines number agreement in SV and VS constructions with singular DPs only, coordinated as in (1) and (2). Moreover, the possibility of singular subject-verb agreement also appears to be sensitive to other properties

of the DPs, such as animacy. In light of Sorace and Keller's (2005) distinctions between hard and soft constraints found in purely syntactic violations vs. syntax-semantics/pragmatics interface violations respectively, Bamyaci et al. (2014) argue that fine-grained distinctions of animacy need be considered for subject-verb agreement in Turkish (for typological observations see Corbett, 2000, 2006; for animacy hierarchy see Haspelmath, 2008). It is not clear whether and how animacy may interact with the Greek subject-verb coordinate agreement structures of the present study. At a first glance, it seems that animate and inanimate DPs allow for singular (partial) agreement with the verb. Consider the example in (7) below:

7.   Xithike   to gala   ke i supa.
        spilled.3s the milk and the soup
        'The milk and the soup spilled over.'

However, given that inanimate subjects are often "derived" as in unaccusative or passive structures, we did not include animacy as a variable in our study. Instead, the subjects used were, all but one, animate.

Processing studies on subject verb agreement have mainly focused on "attraction" errors in which the verb erroneously agrees with an intervening noun bearing number specification different from the head noun of the subject (Franck et al., 2006; Wagers et al., 2009). Findings suggest that such attraction errors in agreement are attested with ungrammatical sentences and are accounted for by a cue-based retrieval mechanism for accessing and comparing previously processed constituents. Tucker et al. (2015) examined agreement errors within the subject constituent in Arabic and found that morphologically discontinuous plural forms need further elaboration for the grammatical features in order for them to be used as processing cues for the retrieval system. These self-paced reading studies focused on adult monolingual data and it is unclear how and whether these attraction errors would affect learners' (child or adult) processing as well.

Child processing studies on subject-verb agreement violations are limited and do not include coordinate subjects. Brandt-Kobele and Höhle (2014) conducted an eye-tracking study with 3 and 5 year old monolingual German speaking children and found that only the older group was sensitive to (un)grammaticality. Preferential listening studies have reported a high sensitivity of monolingual children as young as 2 years old to subject-verb agreement violations (e.g., Soderstrom et al., 2007; Polišenská, 2010; Nazzi et al., 2011. Nevertheless, no online studies are found in the literature that test coordinate subjects in particular.

In Greek, only a limited number of studies (Spyropoulos, 2007; Kazana, 2011) have addressed the syntactic derivation of such constructions and, primarily, from a theoretical perspective. From the processing perspective, it remains unclear how adults and children process these constructions in real-time, whether they are able to rapidly integrate number information and, show sensitivity to the temporary mismatch between plural number in the verb and singular number on the subject in post-verbal coordinate subjects. Finally, it has not been investigated whether

and how monolingual and bilingual speakers of Greek process the (un)grammaticality induced by pre-verbal coordinate subjects and singular number on the verb.

The investigation of coordinate subjects allows us to compare whether children and adults are sensitive to number mismatch between the verb and the subject when processing sentences incrementally, at the point where a mismatch leads to ungrammaticality as soon as the verb is encountered (when the subject is pre-verbal) as opposed to further down in the sentence (when the subject is post-verbal). We anticipate that child data will show the automatic reflex of longer reading times (RTs hereafter) in both cases immediately after the segment in which the mismatch becomes apparent. Sequential bilingual children whose language abilities are usually lower than those of monolingual controls have been shown to be sensitive to SV agreement violations (Chondrogianni and Marinis, 2012) in English sentences with simple subjects despite of their variability in production. The present study investigates whether bilinguals will also be sensitive to subject-verb agreement mismatch in coordinate subjects which are more complex than simple subject DPs.

Finally, given that our bilingual participants are speakers of Albanian and Greek, we considered whether Albanian allows (a) for post-verbal subjects and (b) for partial number agreement with coordinate subjects in post-verbal position. Albanian, like Greek, is a null subject language. As such, post-verbal subjects should be available as a property associated with the null subject parameter (Rizzi, 1986). Although post-verbal subjects are indeed available in Albanian, partial number agreement with post-verbal subjects is accepted by Albanian native speakers but not as strongly as full number agreement (Meniku and Campos, 2016). Moreover, unlike Greek, partial agreement cases are not mentioned in grammar books (Meniku and Campos, 2016). Consider the examples in (8) below (E. Kapia, p.c.):

8.   a. Erdhən   Xhoni dhe Maria.
           Arrived.3p John   and Maria

        b. Erdhi   Xhoni dhe Maria.
           Arrived.3sg. John   and Maria
           'John and Maria arrived.'

Given that L1 and L2 are similar in the relevant respects (post-verbal subjects, full and partial number agreement with post-verbal coordinate subjects), processing data from bilingual Albanian-Greek children should reflect child L2 processing properties rather than (negative) transfer effects.

RESEARCH QUESTIONS AND PREDICTIONS

Our main research question is how coordinate subjects are processed in terms of subject-verb number agreement in VS and SV constructions in Greek by monolingual and bilingual speakers. To this aim, we developed two experiments, an acceptability judgment task (Experiment 1) and a self-paced reading task as the main study (Experiment 2). Adult

monolingual speakers of Greek participated in both experiments so as to establish that the acceptability rates of singular and plural number agreement are indeed sensitive to the position of coordinate subjects (pre-verbal or post-verbal) in the adult grammar and, second, to examine the parsing steps to number resolution. In addition, our study aims to identify whether sequential bilingual children with Albanian as L1 and Greek as L2 process sentences in a way similar to monolingual Greek speaking children in terms of speed and pattern of processing related to SV agreement. This dataset is a valuable addition to the literature of sentence processing in developing grammars (both L1 and L2) and in Greek in particular.

With regard to the adult data, we expect that the availability of partial number agreement in coordinate DPs will be confirmed. In particular, the acceptability data are expected to highlight the difference between pre-verbal and post-verbal coordinate subjects and number agreement options, with post-verbal subjects showing higher tolerance to singular number marking on the verb. Adult processing data are also expected to show sensitivity to the singular-plural number distinction as well as to the singular number option with post-verbal vs. pre-verbal coordinate subjects. However, given the "marked" status of partial agreement discussed above (see Section Introduction), it is possible that adult online data will show a number effect with shorter reading times with plural verbs regardless of subject position.

Turning to the child processing data, we expect that (a) in light of the continuity of parsing hypothesis according to which the structural parser of monolingual children is similar to the adult one (Pinker, 1984; Clahsen and Felser, 2006) monolingual children will show similar processing steps to the adults, and (b) bilingual children will show longer RTs than monolingual children in line with previous sentence processing studies (Marinis, 2007, 2008; Chondrogianni and Marinis, 2012; Chondrogianni et al., 2015). This is partly based on the bilingual children's lower language abilities in their L2 compared to monolingual children. In this study, language ability is measured with an expressive vocabulary test, (see Section Child Participants). In terms of processing patterns for subject-verb number agreement, we expect all participants to show longer RTs in post-verbal subject constructions when the verb is in the plural as opposed to singular because there is a temporary number mismatch between the verb (plural) and the subject (singular) at the first segment following the verb, i.e., the first member of the coordinate subject. In addition, if the derivation which allows singular number marking on the verb with a coordinate subject is different (VP coordination and verb-raising) and more complex than the derivation with plural number marking, we expect a number effect to be attested on the second conjunct or in following segments. If sequential bilingual children process subject-verb agreement qualitatively similarly to monolingual children, the same effect should be attested in both groups of children (for qualitative similarities of bilingual and monolingual children's processing of thematic roles see Marinis and Saddy, 2013). In pre-verbal subject constructions, longer RTs are expected on the verb in singular verb structures compared to plural ones as singular number marking on the

verb is ungrammatical in this context. Recall that in pre-verbal structures, coordination is only allowed as DP-coordination leaving plural number as the only agreement option (Aoun et al., 1994; Spyropoulos, 2007).

MATERIALS AND METHODS

Experiment 1: Acceptability Judgment Task
Participants
Forty adult native speakers of Greek (20 female) were included in Experiment 1. At the time of testing, the mean age was 32 years (age range: 22–60 years old).

Experimental Design
The acceptability judgment task aimed at testing coordinate subject-verb agreement in Greek manipulating two factors: the subject position (pre-/post-verbal) and the number of the verb (singular/plural). The experiment consisted of 96 items; 24 experimental and 72 filler sentences. The experimental items were of similar syllable length and the DPs were definite, singular and animate (with the exception of one inanimate item); half of the DPs involved proper names. The task was conducted as an online survey that lasted approximately 10–15 min. The participants were instructed to evaluate sentences in Greek in a scale of 1–5 with 1 being the score for an unacceptable sentence in Greek and 5 for a fully acceptable one. The conditions are exemplified in (9–12) below:

9. Postverbal Subject, V-singular + Subject
    Emfanistike i Maria ke o Yanis meta tin prosklisi.
    appear$^{-PAST-3SING}$ the^{-NOM} Maria^{-NOM} and the^{-NOM} Yanis^{-NOM} after the^{-ACC} invitation^{-ACC}

10. Postverbal Subject, V-plural + Subject
    Emfanistikan i Maria ke o Yanis meta tin prosklisi.
    appear$^{-PAST-3PLUR}$ the^{-NOM} Maria^{-NOM} and the^{-NOM} Yanis^{-NOM} after the^{-ACC} invitation^{-ACC}

11. Preverbal Subject, Subject + V-singular
    *I Maria ke o Yanis emfanistike meta tin prosklisi.
    the^{-NOM} Maria^{-NOM} and the^{-NOM} Yanis^{-NOM} appear$^{-PAST-3SING}$ after the^{-ACC} invitation^{-ACC}

12. Preverbal Subject, Subject + V-plural
    I Maria ke o Yanis emfanistikan meta tin prosklisi.
    the^{-NOM} Maria^{-NOM} and the^{-NOM} Yanis^{-NOM} appear$^{-PAST-3PLUR}$ after the^{-ACC} invitation^{-ACC}
    'Maria and John turned up after the invitation.'

Out of the 72 filler sentences, half of them were well-formed grammatical sentences (N: 36), as in (5) below, and half ungrammatical (N: 36) as in (13–14) below. Ungrammaticality was always due to violations of inflectional features, such as gender, number or case.

13. I vivliothiki sti sofita ehi pola leromena rafia.
    the^{-NOM} bookcase^{-NOM} in-the^{-ACC} attic^{-ACC} have$^{-PRES-3SING}$ plenty^{-ACC} dirty^{-ACC} shelves^{-ACC}

'The bookcase in the attic has a lot dirty shelves.'

14. *I proti katiki tis neas ipirou efaye rizes
the^{-NOM} first^{-NOM} inhabitants^{-NOM} the^{-GEN} new^{-GEN} continent^{-GEN} eat$^{-PAST\text{-}3SING}$ roots^{-ACC}
'The first inhabitants of the new continent ate roots.'

The Experimental materials were divided into 4 lists in a Latin Square design and fillers were identical in all lists.

Results

To analyze the acceptability data we performed repeated measures analysis of variance (ANOVA) with Number (singular vs. plural) and Subject Position (pre-verbal, post-verbal) as the within subjects variables. **Figure 1** shows the results of the acceptability judgment.

The analysis showed a main effect of Number [$F_{(1, 239)} = 571.069$, $p < 0.001$, $\eta_p^2 = 0.705$], a main effect of Subject Position [$F_{(1, 239)} = 6.052$, $p = 0.015$, $\eta_p^2 = 0.025$], and an interaction between Number and Subject Position [$F_{(1, 239)} = 92.518$, $p < 0.001$, $\eta_p^2 = 0.279$; see **Figure 1**]. Within both types of structures plural receives higher acceptability scores than singular [Post-verbal: $F_{(1, 239)} = 197.037$, $p < 0.001$, $\eta_p^2 = 0.452$; Pre-verbal: $F_{(1, 239)} = 639.387$, $p < 0.001$, $\eta_p^2 = 0.728$]. However, when examining the differences between the two types of structures the comparisons show that singular is more acceptable with post-verbal subjects [$F_{(1, 239)} = 72.419$, $p < 0.001$, $\eta_p^2 = 0.233$], while plural with pre-verbal ones [$F_{(1, 239)} = 26.556$, $p < 0.001$, $\eta_p^2 = 0.100$].

The results of the acceptability judgment task establish the difference in the acceptance rates of singular VPs when their subject precedes or follows them. The online study will identify the processing steps building up the interpretation of those constructions in the adult and child data.

Experiment 2: Self-Paced Reading Task
Adult Participants

One hundred and forty one adult native speakers of Greek (102 female) were included in the main study. At the time of testing, the mean age was 24 years (age range: 18–59 years old). None of those participants completed the acceptability judgment task.

Child Participants

Thirty Greek-Albanian sequential bilingual children (11 female) and forty one monolingual Greek children (33 girls) participated in this study. At the time of testing, the mean age of the bilingual group was 11;3 (age range: 10;3–12;7, standard deviation (SD): 0;6) and the mean age of the monolingual group was 11;2 years of age (age range: 10;2–12;2, SD: 0;5) . There was no significant difference in age between the two groups [$F_{(1, 69)} = 0.101$, $p = 0.752$, $\eta_p^2 = 0.062$]. All participants in the study were typically developing without any history of speech and/or language disorder.

All participants attend monolingual state schools where Greek is used as the majority language. To assess the language history and homogeneity of our bilingual group we collected information on our participants' home language practices in preschool years, early (preschool) and current (bi-)literacy skills, and current language preferences for speaking and listening in daily communication, through extensive questionnaires. Specifically, *home language* questions referred to the child's exposure to each and to both languages from birth up to the age of schooling, i.e., around age 6. The *early (bi)literacy* questions asked for information about whether and in which language(s) family members read books to the child. Questions on *current (bi)literacy* asked for information about children's current language preference/use in writing/reading tasks, and, more specifically, (a) whether the children took language classes in Albanian (L1) and (b) which language was their preferred one for daily writing/ reading tasks (writing lists/letters/cards, reading aloud, texting, emailing, visiting websites, video-gaming, book/magazine reading). Finally, the *current language use* questions asked for the child's language preference/use in oral tasks such as the child's preferred language for oral interaction with family members/friends, for memorizing phone numbers, telling the time, mental counting/calculating and for watching TV/movies. Their answers were used to generate four composite input scores for (a) Greek, (b) Albanian and (c) both languages options.

The children's lexical abilities in Greek were assessed in both languages. A standardized expressive vocabulary test was used for Greek (Vogindroukas et al., 2009, adaptation from Renfrew) and an adaptation of the same task was used for Albanian (Kapia and Kananaj, 2013). These tests provided us with independent measures of our participants' language proficiency in their L1 and L2. To examine whether our bilingual participants formed a homogenous group we examined the factorability of the input factors extracted from the questionnaires and their vocabulary development in each language.

The factorability of all input factors was examined to determine the personal characteristics of the bilingual participants that might further influence their responses. A Principal Axis Factor (PAF) with a Varimax (orthogonal) rotation was conducted on the bilinguals' input profiles. The Kaiser-Meyer-Olkin measure of sampling adequacy was 0.56, close to the recommended value of 0.6, and Bartlett's test of
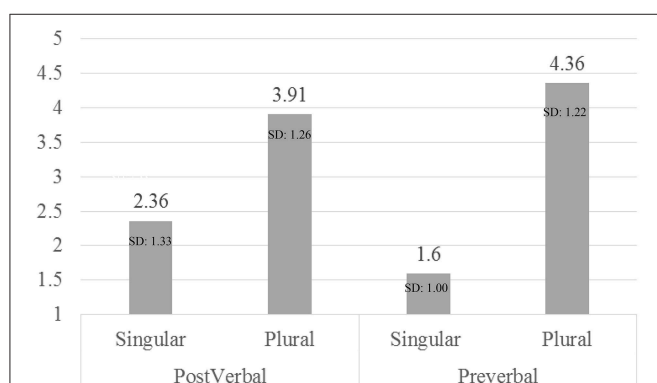
FIGURE 1 | Acceptability judgment task—number and subject position.

TABLE 1 | Descriptive statistics for the factorable bilingual profile characteristics (N: 30).

	Mean (SD)	Skewness (Std. Error)	Kurtosis (Std. Error)
Greek Vocabulary	73% (14)	−0.86 (0.42)	0.07 (0.83)
Albanian Vocabulary	65% (18)	−0.23 (0.42)	−0.59 (0.83)
Greek-dominant Home	31% (18)	0.85 (0.42)	1.08 (0.83)
Albanian-dominant Home	37% (22)	0.49 (0.42)	0.45 (0.83)
Bilingual Home	31% (19)	−0.003 (42)	−0.96 (0.83)

sphericity was significant [$\chi^2_{(136)} = 779.129$, $p < 0.001$] and only loadings >0.30 were considered relevant. The factor analysis showed that 30% of the variance of the data set is explained by the development of Greek lexical abilities and 23% of the variance by the Albanian vocabulary development. Both Greek and Albanian vocabulary scores were very close to normally distributed (see **Table 1**). Out of the questionnaire questions only the home language practices appear to explain some of the variance, with Greek-dominant home practices explaining 9%, Albanian-dominant 9%, and bilingual home ones 8% of the total variance.

Given the outcome of the factor analysis with respect to the role of vocabulary skills in each language and given that the experimental study is a reading task in Greek, we divided the children in two groups; those who scored higher than the mean (+1SD) in the Greek vocabulary task (Group A hereafter, N: 19) and those who scored lower that the mean (Group B hereafter, N: 11). The two groups will be considered in relation to their performance on the self-paced reading task. It is noteworthy that the bilinguals' scores on Greek vocabulary is equivalent to that of 8-year-old monolingual children, indicating at least a 2-year gap in lexical development compared to monolingual controls.

Experimental Design

A self-paced reading task[1] was used to investigate how participants process coordinate subject-verb agreement in Greek. The task manipulated two factors: Subject Position (pre-/postverbal) and Number marking on the verb (singular/plural). The experiment consisted of 106 items; 10 practice sentences, 24 experimental sentences and 72 filler sentences. All experimental and filler items were identical to those of Experiment 1 (Acceptability Judgment Task). Participants controlled the speed of reading each segment by pressing a button on the keyboard. The button press recorded the participants' reading times (RT) per segment. Sentences were segmented in six reading areas: the "Verb," "Subject," "And," "Subject," "PP"/"AdvP" (split in two segments) as in (15–18) below. Slashes indicate segments. Each segment appeared in the middle of the screen and was replaced by the following segment after the participant pressed the spacebar. The last segment appeared with a full stop.

15. V-singular + Postverbal Subject
     Emfanistike / i Maria / ke / o Yanis / meta / tin prosklisi.

---
[1]Stimuli were presented using E-Prime 2.0 software (Psychology Software Tools, 2012. http://www.pstnet.com/).

appear$^{-PAST-3SING}$ / the^{-NOM} Maria^{-NOM} / and / the^{-NOM} Yanis^{-NOM} / after / the^{-ACC} invitation^{-ACC}

16. V-plural + Postverbal Subject
     Emfanistikan / i Maria / ke / o Yanis / meta / tin prosklisi.
     appear$^{-PAST-3PLUR}$ / the^{-NOM} Maria^{-NOM} / and / the^{-NOM} Yanis^{-NOM} / after / the^{-ACC} invitation^{-ACC}

17. Preverbal Subject + + V-singular
     *I Maria / ke / o Yanis / emfanistike / meta / tin prosklisi.
     the^{-NOM} Maria^{-NOM} / and / the^{-NOM} Yanis^{-NOM} / appear$^{-PAST-3SING}$ / after / the^{-ACC} invitation^{-ACC}

18. Preverbal Subject + V-plural
     I Maria / ke / o Yanis / emfanistikan / meta / tin prosklisi.
     the^{-NOM} Maria^{-NOM} / and / the^{-NOM} Yanis^{-NOM} / appear$^{-PAST-3PLUR}$ / after / the^{-ACC} invitation^{-ACC}
     'Maria and John turned up after the invitation.'

As with Experiment 1, out of the 72 filler sentences half were well-formed (N: 36) as in (19) and half ungrammatical (N: 36) as in (20) below. Ungrammaticality in filler items was due to inflectional features such as gender, number or case. Segments are presented in (19) and (20):

19. I vivliothiki / sti sofita / ehi / pola / leromena / rafia.
     the^{-NOM} bookcase^{-NOM} in-the^{-ACC} attic^{-ACC} have$^{-PRES-3SING}$ plenty^{-ACC} dirty^{-ACC} shelves^{-ACC}
     'The bookcase in the attic has a lot dirty shelves.'

20. *I proti / katiki / tis neas / ipirou / efaye / rizes.
     the^{-NOM} first^{-NOM} inhabitants^{-NOM} the^{-GEN} new^{-GEN} continent^{-GEN} eat$^{-PAST-3SING}$ roots^{-ACC}
     'The first inhabitants of the new continent ate roots.'

Yes-no comprehension questions were included for 30% of the total number of sentences to ensure that participants were reading for comprehension. Each question appeared on the screen and participants had to indicate whether the answer was "yes" or "no" by pressing one of the two pre-specified buttons on the keyboard. As with Experiment 1, four lists were created using a Latin Square design. Fillers were identical in all lists. The experiment lasted approximately 15 min.

Results: Adult Data

The responses on the comprehension questions were used to ensure that participants attended to the content of the sentences. A minimum of 90% accuracy on the comprehension questions established that participants were attending and no participant had to be eliminated from further analysis. The variables considered were number on the verb and subject position, i.e., whether the subject appeared pre-verbally or post-verbally. The data obtained included reading times (RTs) on each segment. RTs were screened for extreme values and outliers. Outliers were defined as RTs above or below 2 standard deviations from the mean RT in each condition separately per subject and item. Outliers were replaced with the mean RT for each condition per subject and item once this value was removed. Extreme values

and outliers comprised 2.7% of the adult data (564 out of 20304 data points). Post-verbal and pre-verbal structures were analyzed separately because segments included different words due to the word-order difference. In each data set (post-verbal and pre-verbal structures) we performed repeated measures analysis of variance (ANOVA) with Number (singular vs. plural) as the within subjects factor.

The analysis of post-verbal structures (**Table 2**) showed a main effect of number on the 2nd, 5th, and 6th segments. Specifically, the per subject analysis showed that the participants processed the segment immediately after the verb significantly faster in the singular compared to the plural condition [2nd Segment: $F_{1(1,\ 140)} = 4.992$, $p = 0.027$, $\eta_p^2 = 0.035$] but the singular condition was processed significantly slower than the plural condition the last two sentential segments [5th Segment: $F_{1(1,\ 140)} = 5.477$, $p = 0.021$, $\eta_p^2 = 0.038$; 6th Segment: $F_{1(1,\ 140)} = 14.566$, $p < 0.001$, $\eta_p^2 = 0.095$]. The per item analysis verified the number effect only on the final segment with the plural condition being processed significantly faster than the singular condition [$F_{2(1,\ 23)} = 6.819$, $p = 0.016$, $\eta_p^2 = 0.229$].

The analysis of pre-verbal structures (**Table 3**) showed a main effect of Number only on the 5th segment with shorter RTs in the plural compared to the singular condition, similarly to the post-verbal findings [$F_{1(1,\ 140)} = 18.924$, $p < 0.001$, $\eta_p^2 = 0.120$; $F_{2(1,\ 23)} = 6.955$, $p = 0.015$, $\eta_p^2 = 0.232$].

Results: Child Data

Responses on the comprehension questions were used to ensure participants' attention; both bilingual and monolingual children had a minimum of 80% accuracy in those questions and thus no participant was eliminated from further analysis. As with the adult data, the variable examined was Number on the verb with coordinate subjects appearing either post-verbally or pre-verbally. The data obtained included RTs on each segment. RTs were screened for extreme values and outliers. Extreme values (over 10 s) were identified for each condition separately per subject and item and were removed, leading to the removal of four instances. Extreme values and outliers comprised 0.7% of the bilingual data (29 out of 4320 data points) and 0.9% of the monolingual data (56 out of 5904 data points). In each structure we performed repeated measures analysis of variance (ANOVA) with Number (singular vs. plural) as the within subjects factor and Group (bilinguals vs. monolinguals) as the between subjects factor.

The analysis of post-verbal structures (**Table 4**) revealed a main effect of Group across all segments suggesting overall longer RTs in bilinguals compared to monolingual children [1st Segment: $F_{1(1,\ 70)} = 13.951$, $p < 0.001$, $\eta_p^2 = 0.168$; $F_{2(1,\ 47)} =$

21.610, $p < 0.001$, $\eta_p^2 = 0.320$; 2nd Segment: $F_{1(1,\ 70)} = 22.996$, $p < 0.001$, $\eta_p^2 = 0.250$; $F_{2(1,\ 47)} = 52.276$, $p < 0.001$, $\eta_p^2 = 0.532$; 3rd Segment: $F_{1(1,\ 70)} = 13.802$, $p < 0.001$, $\eta_p^2 = 0.167$; $F_{2(1,\ 47)} = 37.937$, $p < 0.001$, $\eta_p^2 = 0.452$; 4th Segment: $F_{1(1,\ 70)} = 16.254$, $p < 0.001$, $\eta_p^2 = 0.191$; $F_{2(1,\ 47)} = 24.536$, $p < 0.001$, $\eta_p^2 = 0.348$; 5th Segment: $F_{1(1,\ 70)} = 15.973$, $p < 0.001$, $\eta_p^2 = 0.188$; $F_{2(1,\ 47)} = 30.448$, $p < 0.001$, $\eta_p^2 = 0.398$; 6th Segment: $F_{1(1,\ 69)} = 17.479$, $p < 0.001$, $\eta_p^2 = 0.202$; $F_{2(1,\ 47)} = 18.984$, $p < 0.001$, $\eta_p^2 = 0.291$]. Moreover, a main effect of Number on Segment 2, i.e., the first DP immediately after the verb, was found. Specifically, RTs on the first DP were significantly longer when the verb was in the plural than in the singular [2nd Segment: $F_{1(1,\ 70)} = 13.729$, $p < 0.001$, $\eta_p^2 = 0.166$; $F_{2(1,\ 46)} = 13.055$, $p < 0.001$, $\eta_p^2 = 0.321$]. An interaction of Group by Number was only found on the last segment [6th Segment: $F_{1(1,\ 70)} = 4.402$, $p = 0.040$, $\eta_p^2 = 0.060$; $F_{2(1,\ 46)} = 4.124$, $p = 0.048$, $\eta_p^2 = 0.082$] with bilingual children showing longer RTs in plural compared to the singular VPs ($p < 0.001$) and monolingual children longer RTs with singular compared to plural VPs ($p < 0.001$).

In the pre-verbal subject condition (**Table 5**), a main effect of Group across all segments was also found due to the longer RTs in bilingual compared to monolingual children [1st Segment: $F_{1(1,\ 70)} = 10.207$, $p = 0.002$, $\eta_p^2 = 0.129$; $F_{2(1,\ 47)} = 12.883$, $p < 0.001$, $\eta_p^2 = 0.219$; 2nd Segment: $F_{1(1,\ 70)} = 11.607$, $p = 0.001$, $\eta_p^2 = 0.144$; $F_{2(1,\ 47)} = 34.516$, $p < 0.001$, $\eta_p^2 = 0.429$; 3rd Segment: $F_{1(1,\ 70)} = 17.160$, $p < 0.001$, $\eta_p^2 = 0.199$; $F_{2(1,\ 47)} = 24.691$, $p < 0.001$, $\eta_p^2 = 0.349$; 4th Segment: $F_{1(1,\ 70)} = 12.636$, $p = 0.001$, $\eta_p^2 = 0.155$; $F_{2(1,\ 47)} = 24.794$, $p < 0.001$, $\eta_p^2 = 0.350$; 5th Segment: $F_{1(1,\ 70)} = 9.256$, $p = 0.003$, $\eta_p^2 = 0.118$; $F_{2(1,\ 47)} = 15.728$, $p < 0.001$, $\eta_p^2 = 0.255$; 6th Segment: $F_{1(1,\ 70)} = 15.694$, $p < 0.001$, $\eta_p^2 = 0.185$; $F_{2(1,\ 47)} = 16.365$, $p < 0.001$, $\eta_p^2 = 0.262$]. Moreover, a main effect of Number on the segment immediately after the verb was revealed: longer RTs were found after singular verbs compared to RTs for segments following plural verbs [5th Segment: $F_{1(1,\ 70)} = 7.051$, $p = 0.010$, $\eta_p^2 = 0.093$; $F_{2(1,\ 47)} = 7.720$, $p = 0.008$, $\eta_p^2 = 0.144$]. Lastly, no interaction of Group by Number was found suggesting that bilingual and monolingual children process pre-verbal structures similarly.

Lastly, we tested the interaction of the key factorable characteristic of our bilinguals namely Greek vocabulary scores [Group A (high) vs. Group B (low)], with Group as the between subjects factor and Number as the within subjects factor. Both in the post-verbal and pre-verbal conditions no interaction was detected ($p > 0.05$), suggesting that their vocabulary skills did not affect their syntactic processing of coordinate subjects.

TABLE 2 | Adult reading times (in milliseconds) per segment in the postverbal subject condition (SDs in parentheses).

Segments	1st verb	2nd subject	3rd conjunct	4th subject	5th other	6th other
Verb singular	722 (321)	554 (193)	478 (145)	552 (200)	514 (166)	698 (353)
Verb plural	734 (318)	568 (201)	473 (131)	551 (211)	500 (146)	646 (291)

TABLE 3 | Adult reading times (in milliseconds) per segment in preverbal subject condition (SDs in parentheses).

Segments	1st subject	2nd coordinate	3rd subject	4th verb	5th other	6th other
Verb singular	712 (290)	467 (136)	511 (177)	560 (207)	525 (156)	675 (306)
Verb plural	712 (294)	476 (138)	526 (187)	565 (210)	502 (139)	666 (321)

TABLE 4 | Child reading times (in milliseconds) per segment in postverbal subject condition (SDs in parentheses).

Segments		1st verb	2nd subject	3rd conjunct	4th subject	5th other	6th other
Bilinguals	Singular	1543 (1082)	1237 (690)	823 (360)	1281 (926)	917 (509)	1104 (616)
	Plural	1455 (830)	1360 (707)	845 (343)	1306 (795)	898 (396)	1178 (704)
Monolinguals	Singular	1087 (627)	831 (407)	653 (229)	879 (602)	708 (265)	850 (459)
	Plural	1100 (595)	942 (474)	688 (300)	874 (472)	705 (234)	799 (371)

TABLE 5 | Child reading times (in milliseconds) per segment in preverbal subject condition (SDs in parentheses).

Segments		1st subject	2nd conjunct	3rd subject	4th verb	5th other	6th other
Bilinguals	Singular	1476 (1083)	830 (641)	1246 (721)	1201 (679)	868 (343)	1197 (703)
	Plural	1349 (767)	794 (323)	1262 (991)	1234 (669)	804 (293)	1096 (626)
Monolinguals	Singular	1095 (862)	663 (247)	817 (410)	895 (484)	722 (263)	850 (438)
	Plural	1040 (530)	660 (209)	872 (521)	928 (470)	696 (225)	846 (408)

DISCUSSION

The present study examined the processing of Number agreement between coordinate subjects consisting of two singular DPs and the verb, in sentences with the coordinate subject being either in pre-verbal or in post-verbal position. The language studied is Greek and the data included monolingual children and adult Greek speakers and sequential bilingual Albanian-Greek children. Since subject verb agreement is sensitive to both hierarchical and linear (adjacency) constraints, the online data would shed light in the relationship between grammar and parser. Specifically, Greek presents a special case for coordinate subject-verb agreement. Verbs are marked for singular and plural number. Number agreement with coordinate subjects is sensitive to the position of the coordinate subject in that while plural number is the only option with pre-verbal subjects, singular is also possible when the coordinate subject is post-verbal (Spyropoulos, 2007; Kazana, 2011). This is an instance of "partial" agreement attested in other languages too (for Arabic see Aoun et al., 1994, for Slovenian see Marušič et al., 2007, for Hindi see Benmamoun, 2000, and for Serbo-Croatian see Bošković, 2009, 2010). In order to confirm that the singular is indeed an acceptable option in adult Greek we presented an acceptability judgment task including all the sentences used in the online self-paced reading task to a group of adult native speakers of Greek. The results confirmed our predictions. Specifically, singular number agreement in sentences with post-verbal coordinate subjects was significantly more acceptable than in sentences with pre-verbal coordinate subjects. Plural agreement on the other hand was acceptable regardless of subject position.

As suggested in Section Introduction, the derivation of partial agreement (singular verb) involves VP-coordination and V-raising (Aoun et al., 1994; Munn, 1999; Spyropoulos, 2007). In contrast, full agreement assumes DP-coordination and no movement dependency formed. In terms of processing cost, we thus expect that partial agreement would be more complex than full agreement not only because the derivation requires more steps but also because full agreement maps directly onto semantic number agreement while partial agreement does not. In addition, partial agreement is only available with post-verbal coordinate subjects while full agreement is available in all contexts. This restriction adds to the markedness of partial agreement and the associated increased complexity. On these grounds, we predicted that plural agreement would be preferred in online processing showing a number effect at least in the last segments of the sentence with both pre-verbal and post-verbal coordinate subjects. The preference for plural number agreement in both contexts was expected to be found in all groups, although adults were expected to be faster than children, and monolingual children faster than bilinguals. Bilingual children were also expected to show a stronger number effect in the post-verbal condition than monolingual children given the more marginal status of partial agreement with coordinate post-verbal subjects in Albanian (Meniku and Campos, 2016). A number effect was also expected to be found in all groups in the first DP appearing after the plural verb in the post-verbal coordinate subject condition given the local number mismatch. Finally, in the case of pre-verbal subject structures, delays were expected on the singular verb since the coordinate subjects have already been presented and ungrammaticality would be detected on the singular verb itself. This effect should be visible in the

performance of both monolingual and bilingual children as well as in monolingual adult data.

Our results showed that the overall sentence processing patterns of all groups was similar with number effects being found in all groups in a similar way: plural number was processed faster in final segments than singular in both the pre-verbal and the post-verbal subject condition. The monolingual child data appear to support the continuity of parsing hypothesis (Pinker, 1984; Clahsen and Felser, 2006) since the parsing of monolingual children was similar to the adult one. Differences were found in terms of speed of processing; bilingual children were significantly slower compared to monolingual children. In the pre-verbal condition, monolingual and bilingual children performed similarly showing a number effect indicating that they detected ungrammaticality. The data showed a number effect both in the post-verbal and pre-verbal conditions on the segments following the verb; specifically, in post-verbal constructions there was a main effect of number on the first coordinated subject immediately after the verb with plural verbs delaying significantly the processing, and in pre-verbal constructions on the segment following the verb with singular significantly delaying the processing. As anticipated, monolingual and bilingual children did not differ from each other in the pre-verbal condition but we did find an interaction of group by number on the last sentential segment in the post-verbal condition with bilingual children showing slower processing with plural VPs and monolingual children with singular VPs. Monolingual children showed a number effect with faster processing for plural verb structures with both pre-verbal and post-verbal coordinate subjects. We take this effect to indicate that for monolingual children, DP-coordination is used for coordinate subject processing regardless of the subject position. On the other hand, we interpret bilingual children's slower processing of plural verb structures with post-verbal coordinate subjects as an indication of a reanalysis difficulty. In this condition, they encounter a (plural) verb as the first segment of the sentence followed by a singular DP that would be ungrammatical if this was the subject of the sentence. At this point, both monolingual and bilingual children show longer RTs in this condition compared to the condition in which the verb and the first DP are in the singular. The difference between the two groups is however found at the end of the sentence. Monolingual children show shorter RTs in plural compared to singular conditions, a pattern that is similar to adults and demonstrates that they have integrated the two DPs in the coordinated subject construction into a single subject DP and have matched the plurality of the subject with the verb in the plural. Therefore, the condition with a verb in the singular shows elevated RTs. The bilingual children, on the other hand, still have elevated RTs for the condition, in which there was an initial mismatch in number between the plural verb and the first DP in the singular. This could be argued to indicate a difficulty in the reanalysis of their initial parse (first DP is the subject) and integrate the two coordinated DPs as the subject of the verb (for similar effects on the processing of passives see Marinis and

Saddy, 2013). Finally, in the pre-verbal subject condition that tests grammaticality, the number effect is found in both groups in the expected direction.

In conclusion, our study supports findings from other online studies (Marinis, 2007, 2008; Chondrogianni and Marinis, 2012; Chondrogianni et al., 2015) suggesting that bilingual children are slower in incremental processing but not qualitatively different from monolingual children in the grammaticality condition. In this respect, our findings showing a similar number effect in pre-verbal coordinate subjects in bilingual and monolingual children suggest that (un)grammaticality is detected in a similar fashion by the two groups. In contrast, post-verbal structures with coordinate subjects are similar in the two groups only with respect to the delay effect on the first conjunct after the plural verb. This demonstrates again that both groups are sensitive to grammaticality effects in subject-verb agreement constructions. In the final segment, the fact that the two groups show a number effect in the opposite direction is interpreted as a reanalysis and integration problem shown by bilingual children only. Monolingual children show similar processing preferences for plural verbs with coordinate subjects regardless of the subject position. This finding could be interpreted as a VP-coordination and V-raising option being more costly and further delayed in development than the DP-coordination option. Further research into online processing and acceptability judgments of coordinate subjects involving singular and plural DPs as well as pronoun coordination in pre- and post-verbal subject position is required to shed light on the status of the two coordination options.

AUTHOR CONTRIBUTIONS

MK contributed 40% with the setup of the experiment, the data collection and the data analysis. IT contributed 20% as the PI of the research project in which this research is embedded and with the theoretical contribution to the phenomenon investigated, the design of the experiment and the interpretation of the data. TM contributed to the data presentation and write-up and the interpretation of the data. MS contributed to the design of the critical sentences and the theoretical background of the research question.

ACKNOWLEDGMENTS

REFERENCES

Aoun, J., Benmamoun, E., and Sportiche, D. (1994). Agreement, word order, and conjunction in some varieties of Arabic. *Ling. Inquiry* 25, 195–220.

Bamyaci, E., Häussler, J., and Kabak, B. (2014). The interaction of animacy and number agreement: an experimental investigation. *Lingua* 148, 254–277. doi: 10.1016/j.lingua.2014.06.005

Benmamoun, E. (1996). "Agreement asymmetries and the PF interface," in *SOAS Workpaper Papers on Linguistics, Vol. 6* (London), 106–128.

Benmamoun, E. (2000). *The Feature Structure of Functional Categories*. Oxford: Oxford University Press.

Benmamoun, E., Bhatia, A., and Polinsky, M. (2009). Closest conjunct agreement in head final languages. *Ling. Variat. Yearb.* 9, 67–88. doi: 10.1075/livy.9.02ben

Bobaljik, J. D. (2008). "Where's phi? Agreement as a post-syntactic operation," in *Phi-Theory: Phi Features Across Interfaces* and *Modules,* eds D. Harbour, D. Adger, and S. Béjar (Oxford: Oxford University Press), 295–328.

Bošković, Ž. (2009). Unifying first and last conjunct agreement. *Nat. Lang. Ling. Theory* 27, 455–496. doi: 10.1007/s11049-009-9072-6

Bošković, Ž. (2010). "Conjunct-sensitive agreement: Serbo-Croatian vs. Russian," in *Proceedings of FDSL 7.5*, eds G. Zybatow, P. Dudchuk, S. Minor and E. Pschehotskaya (Frankfurt am Main: P.Lang), 31–48.

Bhatt, R., and Walkow, M. (2013). Locating agreement in grammar: an argument from agreement in conjunctions. *Nat. Lang. Linguist. Theory* 31, 951–1013. doi: 10.1007/s11049-013-9203-y

Brandt-Kobele, O.-C., and Höhle, B. (2014). The detection of subject–verb agreement violations by German-speaking children: an eye-tracking study. *Lingua* 144, 7–20. doi: 10.1016/j.lingua.2013.12.008

Chomsky, N. (2001). "Derivation by phase," in *Ken Hale, a Life in Language,* ed M. Kenstowicz (Cambridge, MA: MIT Press), 1–52.

Chondrogianni, V., and Marinis, T. (2012). Production and processing asymmetries in the acquisition of tense morphology by sequential bilingual children. *Bilingualism Lang. Cogn.* 15, 5–21. doi: 10.1017/S1366728911000368

Chondrogianni, V., Marinis, T., Edwards, S., and Blom, E. (2015). Production and on-line comprehension of definite articles and clitic pronouns by Greek sequential bilingual children and monolingual children with Specific Language Impairment. *Appl. Psychol.* 36, 1155–1191. doi: 10.1017/S0142716414000101

Clahsen, H., and Felser, C. (2006). Grammatical processing in language learners. *Appl. Psychol.* 27, 3–42. doi: 10.1017/S0142716406060024

Corbett, G. (1991). *Gender*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139166119

Corbett, G. (2000). *Number*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139164344

Corbett, G. (2006). *Agreement*. Cambridge: Cambridge University Press.

Franck, J., Lassi, G., Frauenfelder, U. H., and Rizzi, L. (2006). Agreement and movement: a syntactic analysis of attraction. *Cognition* 101, 173–216. doi: 10.1016/j.cognition.2005.10.003

Harbert, W., and Bahloul, M. (2002). "Postverbal subjects in Arabic and the theory of agreement," in *Themes in Arabic and Hebrew Syntax,* eds J. Ouhalla and U. Shlonsky (Dordrecht: Kluwer), 45–70.

Haspelmath, M. (2008). "Creating economical morphosyntactic patterns in language change," in *Language Universals and Language Change,* ed J. Good (Oxford: Oxford University Press), 185–214.

Holton, D., Mackridge, P., and Philippaki-Warburton, I. (1997). *Greek: A Comprehensive Grammar of the Modern Language*. London: Routledge.

Johannessen, J. B. (1998). *Coordination*. New York, NY: Oxford University Press.

Kapia, E., and Kananaj, A. (2013). The Albanian Vocabulary Instrument. Tiranë: Center for Albanian Studies, Institute of Linguistics and Literature.

Kazana, D. (2011). *Agreement in Modern Greek Coordinate Noun Phrases.* Unpublished Doctoral Dissertation. Colchester: University of Essex.

Marinis, T. (2007). "On-line processing of passives in L1 and L2 children," in *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)*, eds A. Belikova, L. Meroni, and M. Umeda (Somerville, MA: Cascadilla Proceedings Project), 265–276.

Marinis, T. (2008). "On-line processing of sentences involving reflexive and non-reflexive pronouns in L1 and L2 children," in *Language Acquisition and Development. Proceedings of GALA 2007*, eds A. Gavarró and M. J. Freitas (Newcastle upon Tyne: Cambridge Scholars Publishing), 348–358.

Marinis, T., and, D., Saddy (2013). Parsing the passive: comparing children with Specific Language Impairment to sequential bilingual children. *Lang. Acquisit.* 20, 155–179. doi: 10.1080/10489223.2013.766743

Marušič, F., Nevins, A., and Saksida, A. (2007). "Last-conjunct agreement in Slovenian," in *Proceedings of Formal Approaches to Slavic Linguistics: the Toronto Meeting 2006*, eds R. Compton, M. Goledzinowska, and U. Savchenko (Ann Arbor, MI: Michigan Slavic Publications), 210–227.

Meniku, L., and Campos, H. (2016). *Discovering Albanian 2.* Madison: Wisconsin University Press.

Munn, A. (1999). First conjunct agreement: against a clausal analysis. *Ling. Inquiry* 30, 643–686. doi: 10.1162/002438999554246

Nazzi, T., Barrière, I., Goyet, L., Kresh, S., and Legendre, G. (2011). Tracking irregular morphophonological dependencies in natural language: evidence from the acquisition of subject-verb agreement in French. *Cognition* 120, 119–135. doi: 10.1016/j.cognition.2011.03.004

Pinker, S. (1984). *Language Learnability and Language Development*. Cambridge, MA: Harvard University Press.

Polišenská, D. (2010). *Dutch Children's Acquisition of Verbal and Adjectival Inflection*. PhD thesis, LOT.

Rizzi, L. (1986). Null objects in Italian and the theory of pro. *Ling. Inquiry* 17, 501–557.

Soderstrom, M., White, K. S., Conwell, E., and Morgan, J. L. (2007). Receptive grammatical knowledge of familiar content words and inflection in 16-month-olds. *Infancy* 12, 1–29. doi: 10.1111/j.1532-7078.2007.tb00231.x

Sorace, A., and Keller, F. (2005). Gradience in linguistic data. *Lingua* 115, 1497–1524. doi: 10.1016/j.lingua.2004.07.002

Spyropoulos, V. (2007). "Some remarks on conjunction and agreement in Greek: implications for the theory of agreement," in *Proceedings of the 7th International Conference of Greek Linguistics* (York, UK), 1–12.

Tucker, M. A., Idrissi, A., and Almeida, D. (2015). Representing number in the real-time processing of agreement: self-paced reading evidence from Arabic. *Front. Psychol.* 6:347. doi: 10.3389/fpsyg.2015.00347

Vogindroukas, I., Protopapas, A., and Sideridis, G. (2009). *Experiment on the Expressive Vocabulary (Greek Version of Renfrew Word Finding Vocabulary Test)*. Chania: Glafki.

Wagers, M. W., Lau, E. F., and Phillips, C. (2009). Agreement attraction in comprehension: representations and processes. *J. Mem. Lang.* 61, 206–237. doi: 10.1016/j.jml.2009.04.002

CrossMark

Linguistic and Cognitive Skills in Sardinian–Italian Bilingual Children

Maria Garraffa[1]*, Madeleine Beveridge[2] and Antonella Sorace[2]

[1] Department of Psychology, School of Life Science, Heriot-Watt University, Edinburgh, UK, [2] School of Philosophy, Psychology and Language Sciences, The University of Edinburgh, Edinburgh, UK

We report the results of a study which tested receptive Italian grammatical competence and general cognitive abilities in bilingual Italian–Sardinian children and age-matched monolingual Italian children attending the first and second year of primary school in the Nuoro province of Sardinia, where Sardinian is still widely spoken. The results show that across age groups the performance of Sardinian–Italian bilingual children is in most cases indistinguishable from that of monolingual Italian children, in terms of both Italian language skills and general cognitive abilities. However, where there are differences, these emerge gradually over time and are mostly in favor of bilingual children.

Keywords: minority languages, grammar, bilingualism, executive functions, Sardinian, object relatives

INTRODUCTION

Multilingualism is the norm in many parts of the world: according to some conservative estimates (Tucker, 1998), at least half of the world's population speaks two or more languages. While many factors contribute to the increase in bilingualism in Europe, including transnational population mobility and the status of English as a *lingua franca*, bilingualism in regional minority languages is declining due to the lack of intergenerational transmission (see Romaine, 2007; Extra and Gorter, 2008). Fewer parents speak minority languages to their children because of their perceived lack of 'usefulness' and other more general misconceptions on early bilingualism. A similar gap is seen in research into different types of bilingualism. Bilingualism is the object of much linguistic and cognitive research that investigates different aspects of development and use, but bilingualism involving minority languages has not received the same attention as bilingualism involving prestigious languages with wide currency. This paper makes a contribution to redressing the balance by presenting the results of a pilot study on the linguistic and cognitive abilities of children who speak Sardinian as a minority language and Italian as the majority language. We will first briefly summarize research on language development in bilingualism, with an emphasis on grammatical models and general cognition. This will be followed by some notes on the status of Sardinian as a minority language. We will then present the methods employed in the collection of data and the results of statistical analyses. Finally, the data will be discussed against the wider context of bilingualism in regional minority languages.

Language and Cognition in Bilingual Children: Highlights of Previous Research
Morphosyntactic Development
The central question underlying research on bilingual syntactic acquisition is whether bilingual children differentiate their two languages at all stages of development, and whether the two language grammars influence each other. In spite of consensus in early research that bilingual first

language acquisition is characterized by independent and parallel acquisition of syntax (Meisel, 1989; De Houwer, 1990; Genesee et al., 1995), more recent research has revealed a more nuanced picture.

For example, Dopke (1998) and Yip and Matthews (2007) reported cross-linguistic effects of one language on the other at the syntactic level, from the dominant language, or the language of the environment, to the weaker language. The effects of dominance and of the amount of input in the weaker language are solidly attested. Bernardini and Schlyter (2004) found syntactic effects of Swedish on Italian and French in Swedish-dominant bilinguals; Meisel (2007) found a lower mean length of utterance (MLU) but no divergent syntactic patterns in the weaker French of French–German bilinguals; Gathercole (2007) reported that monolingual English children outperform school-age English–Spanish bilinguals who are dominant in Spanish in measures of both mass/count distinction and gender; Paradis et al. (2011) studied regular and irregular English past tense in English-dominant and French-dominant children, reporting that English-dominant children scored lower than monolinguals only for irregular forms, but French-dominant children scored lower on both English regular and irregular forms. A similar but more qualified conclusion was reached by Blom (2010) who showed clear input effects in younger Dutch–Turkish bilinguals: Turkish-dominant children were delayed in acquiring the relationship between finiteness and subject realization in Dutch, but Dutch-dominant children were not. Blom argued that reduced input quantity does slow down grammatical development. However, these differences are limited to the weaker language of bilingual children, and are visible only in situations of clearly reduced input. When bilingual children receive balanced input in the two languages, other factors such as age of first exposure and consistent input for a particular structure play an important role. Unsworth et al. (2014), for example, showed that highly regular and consistent grammatical gender in Greek is acquired in similar ways by simultaneous English–Greek bilinguals and monolingual Greeks, but the similarity breaks down in consecutive older bilingual children. In contrast, the inconsistent system of gender in Dutch is acquired late both by monolingual Dutch children and by English–Dutch bilinguals, regardless of age of first exposure.

Cross-linguistic effects in bilingual development may be selective and asymmetric for other reasons. Müller and Hulk's (2001) seminal work argued that structures at the interface between morphosyntax and discourse are vulnerable to cross-linguistic influence in early bilingual language development, but core syntactic structures are not. Subsequent research refined this hypothesis. For one thing, it was shown that not all structures that satisfy the 'interface' requirements show evidence of cross-linguistic influence (see Unsworth, 2005 on optional infinitives in English–German bilinguals). Furthermore, phenomena at the syntax-pragmatics interface, such as the interpretation of pronominal anaphoric forms, take longer to be acquired than phenomena at the syntax-semantics interface such as the use of determiners in generic vs. specific plural nouns (Paradis and Navarro, 2003; Serratrice et al., 2004; Serratrice, 2007; Sorace and Serratrice, 2009). An emerging striking generalization is that delays and inconsistency at the syntax-pragmatics interface

have been attested in bilingual children regardless of whether the two languages are grammatically similar, and have been found to also characterize late bilinguals in both the L2 and the L1 (Sorace, 2011, 2012). These parallelisms suggest that the reason for the generality of these effects in bilingualism may lie in extra-linguistic general cognitive factors, rather than in language-specific effects of one grammar over the other.

In the study reported in this paper we investigate possible effects of Sardinian on Italian in school-age children who grow up in an environment where Italian is the majority language, but who are exposed to proportionally more Sardinian in early childhood until they start schooling. We chose to focus on comprehension of a range of productive syntactic structures of Italian with different degrees of complexity, as a first step toward establishing whether there are indeed effects of Sardinian on Italian at the beginning of the schooling process and whether these effects decrease with more exposure to Italian. The structures tested were active and passive structures, coordination, dative structures, topicalisation/left dislocation, and subject and object relatives (see **Tables 1** and **2** and see Studies of Regional Minority Languages).

Cognitive Effects of Bilingualism

Recent research on bilingualism has revealed that the bilingual experience can have effects on general cognition beyond the language domain (see Bialystok, 2009; Baum and Titone, 2014; Costa and Sebastian-Galles, 2014 for overviews). The most consistent empirical finding is that of advantage in attentional aspects of executive functions. Adopting Miyake and Friedman's (2012) tripartite distinction of executive functions into updating, shifting, and inhibition, one can say that the jury is still out as to precisely which component(s) are affected by bilingualism. What seems to be clear, however, is that some of these effects are greater in bilingual children and older bilingual speakers than in young bilingual adults, possibly because the effects are more visible when executive functions are either developing or declining but are not at their peak (Craik and Bialystok, 2006). In bilingual children, advantages have been found in metalinguistic tasks requiring a focus on form in the presence of a distracting meaning (Bialystok, 1988, 1992). Executive control may be involved in these tasks in order to ignore the meaning and focus on form. Similarly, advantages have been reported for the development of theory of mind (ToM) and pragmatic/conversational abilities (Goetz, 2003; Siegal et al., 2009, 2010), which may involve executive control in the suppression of ones' own perspective when focusing on that of others.

Discussions of the reasons behind the bilingual advantage rely on defining how the two languages are processed in the brain, how they are accessed and how they interact with one another. One theory that has attracted much consensus is the joint activation model (Green, 1998), which assumes that both languages are always active regardless of whether the context of communication is monolingual or bilingual. The bilingual speaker therefore has to suppress the language not in use, or alternatively to enhance activation of the target language (Costa et al., 2006). The core of the debate revolves around whether the main advantage displayed by bilinguals is the ability to focus

on the desired information while 'ignoring' (but not 'inhibiting') the distracting information, or whether it crucially lies instead with the ability to *inhibit* irrelevant information or distracters (Bialystok, 2009). While Bialystok (2009) puts more weight on inhibitory control as the key force in the language selection process, she recognizes that one mechanism is not necessarily mutually exclusive of the other: it could be the case that both inhibiting and ignoring can allow the bilingual speaker to use one language without interference from the other (see also Adaptive Control Hypothesis; Green and Abutalebi, 2013). Depending on the type of bilingual experience and how these experiences 'sculpt' the bilingual brain, one might expect to see different effects on general cognitive abilities. Bilinguals have been shown in some studies to outperform monolinguals not only in trials that require inhibitory control of distracting information, but also in trials where no distracting information is present: this fact suggests that the cognitive abilities affected by bilingualism may be broader and more general than inhibitory control (Hilchey and Klein, 2011). It should be added that a bilingual advantage has also been found in a few studies of infants (see, e.g., Kovács and Mehler, 2009) who do not yet experience language control in production (but see Blumenfeld and Marian, 2011, on how inhibitory control affects comprehension too).

It is possible that different types of bilingual experience may lead to different (or null) effects on cognitive abilities. For instance, Costa et al. (2009) proposed that speakers with highly separated and predictable domains of use for each language – thus with a low level of switching required – may not show advantages. Similarly, Prior and Gollan (2011) suggest that an advantage in task switching may arise only in bilinguals who frequently switch between languages. The presence of bilingualism in all societal contexts may have an effect, as well as the relatedness of language pairs (Costa et al., 2009; see Grohmann, 2014 on 'language proximity' as an important factor for simultaneous child bilingualism). With this in mind, it is important to gather data from different types of bilinguals, with different language backgrounds, to gain a fuller picture of the effects of bilingualism in particular domains.

The most recent debate has centered in particular on the replicability of the 'bilingual advantage,' which a number of studies have failed to find (Paap and Greenberg, 2013; Duñabeitia et al., 2014; Paap, 2014). Some researchers interpret these null results as questioning the validity of previous results showing a bilingual advantage (see de Bruin et al., 2015; Valian, 2015). Others view the failure to replicate in some studies as a normal manifestation of variation due to interactions with poorly understood factors (age at testing, language combination, patterns of bilingual language use, education levels, societal attitudes, etc.), and ultimately as a welcome incentive to carry out more research in different bilingual settings. Bilingualism with regional minority languages, in particular, is a setting that has generated a sparse and inconsistent picture (see below). Furthermore, there is a need for more research that compares child and adult bilinguals in order to trace the developmental trajectory of the effects of bilingualism over the lifespan. More research is also needed to compare children who become bilingual at different stages of childhood (see Bialystok et al.,

2012). The Sardinian context offers a unique opportunity to study the emergence of bilingualism in a minority language and its effects over time in school-age children who receive instruction in the majority language.

Bilingualism in Regional Minority Languages

As a broad group, minority languages tend to differ in significant ways from majority languages with respect to (i) quality and quantity of input, (ii) social status and attitudes toward the language, and (iii) motivation toward bilingualism. First, a significant proportion of languages of the world today are currently facing a drastic decline in numbers of speakers (Nettle, 1999; Crystal, 2000; Grenoble and Whaley, 2006). Thus, the range of different speakers a child acquiring the language has exposure to may be limited. Having exposure to a range of different speakers is important in the acquisition of any language and may affect the child's language proficiency (Houston and Jusczyk, 2000). It can also be the case with minority languages (likely more so than with majority languages) that teachers, parents and others passing on the language to the child may be second language speakers/learners themselves. This situation inevitably generates a different type of exposure for the child learning a minority language, compared with a child learning a majority language and who is likely to have input from a wide range of different, native speakers. Second, the often unstable or turbulent political history of the minority language may negatively affect the linguistic experience of children. This may be manifested, for example, in the form of lack of institutional support toward the language or in parental lack of motivation to speak the language due to its perceived inutility (Crystal, 2000). Sardinian is no exception in this broad picture.

Studies of Regional Minority Languages

The cognitive effects of bilingualism in minority languages have been investigated in a limited number of studies, which provide an inconsistent picture. On the one hand, no bilingual advantage in executive functions was found in studies of Welsh–English bilinguals (Gathercole et al., 2014) and Basque–Spanish bilinguals (Duñabeitia et al., 2014). These studies focused on communities where the minority language has an officially recognized and protected status, yet no differences were reported. On the other hand, other studies do show an advantage for bilingual speakers of minority languages. Antoniou et al. (2014) tested children in Cyprus who were bilingual (or 'bilectal') in Greek and Cypriot Greek, and found that they outperformed age-matched monolingual children on all measures of cognitive control, although not on all vocabulary measures. Lauchlan et al. (2013) compared Gaelic–English and Sardinian–Italian bilingual and monolingual English and Italian children in Scotland and Sardinia on measures of cognitive control, problem-solving ability, metalinguistic awareness, and working memory. The results showed a global bilingual advantage over the monolinguals in two of the four measures used. In addition, the bilingual Scottish children significantly outperformed the bilingual Sardinian children: this difference is interpreted as a consequence of the fact that the bilingual Scottish children

received Gaelic-medium education, in contrast to the Sardinian bilingual children who mostly speak the minority language only at home. Finally, Vangsnes et al. (2015) looked at the effects of bidialectal literacy in the two Norwegian standards Nynorsk (the minority system) and Bokmål (the majority system) in the minority group of pupils who are schooled in Nynorsk. The data show that these pupils perform better than average in national tests of English, reading and arithmetic once socio-economic factors are controlled for.

Sardinian

Most scholars regard Sardinian as a separate Romance language (Harris and Vincent, 1988; Posner, 1996). The long period of independent development following the fall of the Roman Empire distinguishes it clearly from other Romance languages, and it is not intelligible to speakers of Italian. However, the present-day sociolinguistic reality is such that most speakers of Standard Italian probably consider it to be a "dialect" of Italian. Sardinian tends to be used in local and/or informal settings, while Standard Italian is the expected language in official contexts, in cities, in church and in school.

The Sardinian regional government commissioned a comprehensive study of language use in the early part of the 21st century (Oppo, 2007), based on a sample of approximately 2400 respondents aged 15 and above from all over the island. According to this study, nearly 70% of respondents reported that they speak a "local language" (term referring to any local variety of Sardinian, as well as to the other languages spoken by small communities on the island such as Gallurese and Catalan) and nearly 30% said they understood one but did not speak it; only 2.7% claimed no knowledge of a local language. The study also confirmed that there are substantially fewer speakers of local languages in towns and cities with more than 20,000 inhabitants than in villages and rural areas. There are probably no monolingual speakers of Sardinian anywhere on the island, though there are certainly elderly speakers who are more at ease in Sardinian than in Italian.

Oppo's study also briefly reports the results of a similar survey of approximately 270 children under 14. The proportions are markedly different from the adult figures: just over 40% reported speaking a local language; just over 35% said they understood but did not speak a local language; and more than 20% said they neither spoke nor understood a local language. The substantially smaller proportion of children than adults who report using a local language clearly points to the endangered status of Sardinian as a whole. There are still parts of the island, such as the Nuoro province in central Sardinia, where children routinely learn Sardinian in the family before learning Italian at school, but there are many more children who learn Italian in the family and never acquire Sardinian.

Sardinian and Italian: A Brief Comparison

Although the grammars of Sardinian and Italian share a common origin, they are not identical – for a general description of the syntactic differences between the two languages, see Jones (1993) and Bolognesi (2013). One difference that is relevant for the structures in focus here concerns the passive structure, for

which dialectal variation is observed. In particular, the passive is possible but dispreferred by speakers in the central Sardinian areas where the data were collected, whereas speakers from southern regions find it more acceptable, possibly because of the stronger influence of Italian. Other relevant differences are the prepositional marking of direct objects and the clitics doubling with indirect objects, which are common in all varieties of Sardinian but ungrammatical in Italian.

Another point of interest is how bilingual children deal with structures that have been reported to be developmentally late in monolingual acquisition. A well-known example is relative clauses, which have been identified in several studies as difficult to acquire in different languages (see Adani, 2011 for an overview). Object relatives, in particular, develop rather late in monolinguals. A theoretical account of the source of complexity for object relatives originally proposed for adults with acquired language disorders and children with atypical language development (Garraffa and Grillo, 2008; Grillo, 2008; Contemori and Garraffa, 2010) and successfully extended to typical language development (Friedmann et al., 2009), is in terms of the intervention of the lexical subject on the long distance dependency established between the relative head and its original position. This intervention effect is schematically shown below.

DP [DP. . .. < DP >]
. . .il cane che il bambino insegue < il bambino >.
The dog that the child chases

Object headed relative clauses are more difficult to produce and comprehend compared to subject headed relative clauses. Production studies in fact reveal different strategies adopted by monolingual speakers in order to produce simpler sentences not subject to intervention in place of an object relative, but still preserving the meaning of the sentence (Contemori and Belletti, 2012).

One well-attested strategy to avoid intervention is replacing object relative with a passive object relative, POR (i.e., *Il cane che è inseguito dal bambino*, 'the dog that is chased by the child,' in place of *il cane che il bambino insegue*, the dog that the child chases). In order to use the POR strategy productively it is necessary to fully master the passive morphology that is the trigger for the movement of the verb phrase not subject to intervention (see Collins, 2005 for a detailed approach on passives sentences). Another productive strategy to avoid the complexity of the object relative was reported by Adani et al. (2010), where an ameliorative effect on comprehension of ORs was attested in the case of sentence with argument number mismatch (i.e., *Il leone che I coccodrilli stanno toccando è seduto per terra* 'the lion-SG that the crocs-PL are touching is sitting-SG on the floor'). Both the passive structure and verbal inflection strategies required a full command of the morphosyntactic aspects of the language. Adults as well as monolingual children at young ages either did not produce object relatives, replacing them with passive object relatives, or are more likely to produce object relatives when there is a morphological mismatch between the arguments. The question is whether these difficulties would affect bilingual children to the same extent as monolinguals in a comprehension task, given that Sardinian

relative clauses are structurally similar to Italian relative clauses (see **Table 1** below).

Research Questions

This pilot study aims to address these questions:

(a) Do Sardinian–Italian bilingual children have a disadvantage compared with monolingual Italian children in their comprehension abilities of Italian when they start being schooled in Italian? If they do, is the disadvantage manifested only for particular structures? If there is a difference between bilingual and monolingual children, does it change over time due to age and more experience of Italian in the school setting?

(b) Do Sardinian–Italian bilingual children have an advantage compared to monolingual Italian children in general cognitive abilities related to attentional control and executive functions? If there is a difference between bilingual and monolingual children, does it change over time due to age and more experience of Italian in the school setting?

MATERIALS AND METHODS

Participants

Ninety five children from nine primary schools in the towns of Fonni, Orgosolo, Mamoiada, Nuoro, Desulo, Tonara, Bitti, Lula, and Orune, all in the Nuoro Province, participated in the study. All children were attending the first or the second year of primary school, where the language of instruction is Italian. 10 children were excluded because they did not meet standardized criteria in one or more screening background tests (see below). The final sample included 85 children whose ages ranged from 6 to 9 years and 1 month. For the majority of bilingual children, exposure to Italian occurred at school; therefore, the amount of time spent in education was considered an important predictor of Italian competence. At the time of testing, 18 of the bilingual children and 20 of the monolingual children were finishing their first year of Italian primary school; 22 of the bilingual children and 25 of the monolingual children were finishing their second year of Italian primary school. Thus, the children represented four groups: (a) 18 bilinguals with 1 year of Italian schooling, (b) 22 bilinguals with 2 years of Italian schooling, (c) 20 monolinguals with 1 year of Italian schooling, and (d) 25 monolinguals with 2 years of Italian schooling.

Tasks

Background Measures

Parental background questionnaire

Children's language background and exposure to both Italian and Sardinian was measured using an adapted version of the UBILEC, a comprehensive parental questionnaire measuring quantitative and qualitative aspects of language exposure (Unsworth, 2013a; Unsworth et al., 2014). The UBILEC questionnaire captures the amount of target language exposure over time in the past considering possible variation in early language development, such as language use during holiday and languages spoken in

daycare or at school. To better quantify language competence in each language we looked at the information provided for each child by the cumulative language exposure index, which is part of UBILEC: this measured how much input was received from each parent and any other adults over time both at home and outside the home. The cumulative index is a detailed estimation of children's language exposure over the years and a more accurate one compared to the traditional index of exposure that measures the differential amount of exposure between the languages (see Unsworth, 2013b for a detailed review). Children who scored lower than 3.3 on the UBILEC cumulative exposure index parameter for Italian were classified as bilingual. This was calculated as a median cut-off of the score reported for each child. Accordingly, 40 children were classified as bilingual and 45 children as monolingual. The bilingual children spoke Sardinian at home and in the community, and Italian at school. Given that Sardinian is the language commonly spoken in daily interactions in the Nuoro Province (Oppo, 2007), the monolingual children may also have been exposed to some Sardinian in the surrounding community, but Italian is the language spoken in their family as well as in day care or at school.

Raven's Colored Progressive Matrix test

All children completed the Raven's Colored Progressive Matrix (CPM) test of general intelligence (Raven et al., 1998) as an inclusion criterion to exclude any intellectual impairment. Children who performed within 2 SD of the age-corrected standardized score were included in the study.

Peabody Picture Vocabulary Test of receptive vocabulary (PPVT-4)

Recent discussions about the relative size of age-matched monolingual vs. bilingual children's vocabulary (e.g., Bialystok, 2009; Bialystok et al., 2010) raise the possibility of differences in Italian language vocabulary between the monolingual and bilingual groups. The Peabody Picture Vocabulary Test of receptive vocabulary (PPVT-4, Stella et al., 2000) was therefore administered to all the children to establish their receptive Italian vocabulary knowledge. The test is incremental, and a basal score is established when the child makes six errors in eight consecutive responses. All children with a performance within 2 SD of the age normed transformed score were included in the study.

Digit span task

Several accounts suggest that areas of cognitive development (for example, executive function) are facilitated by short term memory (e.g., Gordon and Olson, 1998). Phonological memory was therefore assessed using a digit span test adapted from (Orsini et al., 1987; see Gathercole, 1998 for a review). For inclusion into the study, children had to show a digit span of ≥4 digits. No children were excluded.

Non-word repetition task

Non-word repetition has been shown to be a reliable index of verbal memory development and a clinical marker for detecting language impairment. A number of studies have reported that bilingual children are highly proficient in this task, sometimes showing an advantage over monolingual speakers (Tamburelli

et al., 2015), but Guasti et al. (2013) found no differences between early second language learners and age-matched monolinguals Italian speakers. We therefore tested children on the non-word repetition task developed by Cornoldi et al. (2009) to exclude language impairment in both groups. To be included in the study, children had to achieve a non-word repetition score of at least 10 syllables. No children were excluded.

Test Measures

Receptive grammatical knowledge

Grammatical competence in Italian was measured using the COMPRENDO test (Cecchetto et al., 2012); a picture-matching task assessing sentence comprehension in Italian across syntactic structure types. The types of sentences included are all semantically reversible (with both nouns possibly acting as subject or object of the verb) and span structural complexity over seven conditions, shown in **Table 1**. As Section "Sardinian and Italian: A Brief Comparison," Sardinian is both similar and different from Italian with respect to these structures. This is shown in **Table 1**.

There were three items per condition with a total of 21 items per list, resulting in a 7×3 design. For each sentence, the child was asked to select one of four pictures (see example in **Figure 1**). The correct picture matched the sentence content:

FIGURE 1 | COMPRENDO sample picture for dative target sentence "Il bambino da la torta alia mamma" *(The boy gives the cake to the mother).*

for the sentence "La mamma da la torta al bambino" (*The mum gives the cake to the boy),* the picture showed a mother giving a cake to a young boy. In addition, there were three incorrect "distractor" pictures. The *reversal distractor* depicted the same actors in reversed roles (e.g., a boy giving a cake to

TABLE 1 | Sentence structure types tested in the COMPRENDO receptive test of Italian and translations in Sardinian.

Sentence type	Italian example	Sardinian translation
Active	Il cane morde il gatto *The dog bites the cat*	Su cane mossigat (a) sa gato *The dog bites the cat*
Dative	La mamma dà la torta al bambino *The mother gives the cake to the boy*	Sa mamma li dat su durce a su pitzinneddu *The mother to-him gives the cake to the boy*
Coordinate object	Il bambino insegue il cane e il gatto *The boy chases the dog and the cat*	Su pitzinneddu pressighit su cane e sa gato *The boy chases the dog and the cat*
Passive	Il bambino viene inseguito dal cane *The boy is chased by the dog*	Su pitzinneddu est pressighidu dae su cane *The boy is chased by the dog* Su pitzinneddu lu pressighit su cane *The boy him chases the dog*
Topicalised OSV-number mismatch	La bambina, i nonni la inseguono *The girl, the grandparents chase her*	Sa pitzinnedda la pressighint sos mannois *The girl, the grandparents chase*
Subject relative	Il nonno spinge il cane che morde il gatto *The grandfather pushes the dog that bites the cat*	Su mannoi ispinghet su cane chi mossigat sa gato *The grandfather pushes the dog that bites the cat*
Object relative	La mamma guarda il cane che il bambino insegue *The mother looks at the dog that the boy chases*	Sa mamma abbaidat su cane chi su pitzinneddu pressighit *The mother looks at the dog that the boy chases*

TABLE 2 | Mean age, cumulative length of exposure to Italian, and performance on background tests: RAVEN, PPVT-4, digit span, and non-word repetition tasks across groups (raw scores and SD).

| Group N = 85	Age Years Mean (SD)	UBILEC Cumulative exposure index	RAVEN Mean (SD)	PPVT- 4 Mean (SD)	Digit span Mean (SD)	Non-word repetition Mean (SD)
Bilingual Y1 N = 18	6.65 (0.32)	1.0	24.00 (3.56)	96.83 (7.96)	4.40 (0.60)	13.45 (2.19)
Bilingual Y2 N = 22	7.80 (0.47)	0.8	24.41 (3.9)	102.50 (13.29)	5.10 (0.70)	11.75 (1.48)
Monolingual Y1 N = 20	6.60 (0.31)	4.4	22.80 (3.21)	99.35 (16.11)	4.80 (0.68)	12.95 (2.29)
Monolingual Y2 N = 25	7.68 (0.26)	4.6	24.68 (3.74)	100.24 (11.53)	4.84 (0.61)	12.48 (2.06)
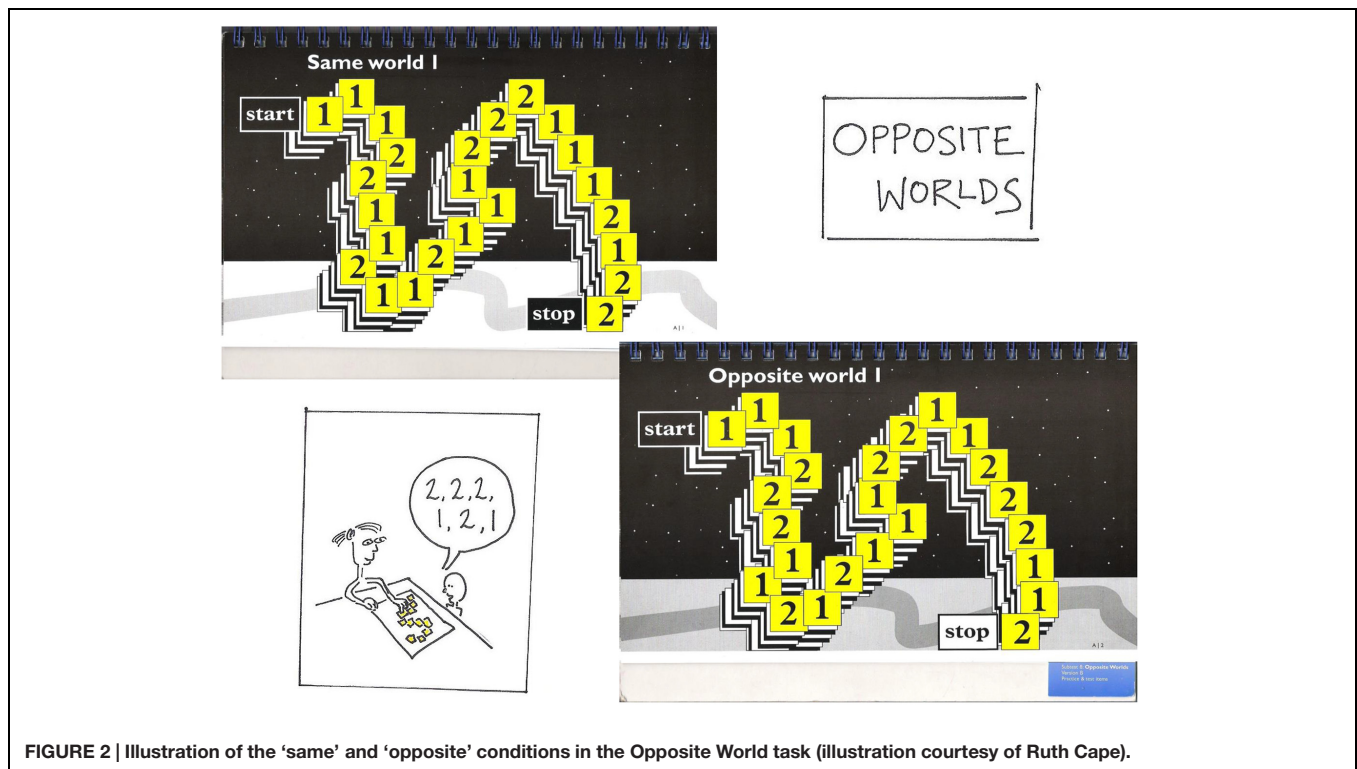
FIGURE 2 | Illustration of the 'same' and 'opposite' conditions in the Opposite World task (illustration courtesy of Ruth Cape).

his mother). The *verbal distractor* depicted the actors in the same thematic roles, but completing a different action (e.g., the mother caressing the boy). The *nominal distractor* kept the same action (e.g., giving), but replaced all the nouns (both the actors and the object; e.g., *The grandmother gives the keys to the girl*).

The task requires children to map the thematic roles (i.e., *Who is doing what to whom?*) in relation to the syntactic form of the sentence. This is a test of grammatical knowledge. However, general cognitive abilities such as executive control might be involved in this task, since competing interpretations have to be held in memory, and the incorrect ones must be inhibited.

Opposite world task

This task is part of the Test for Everyday Attention for children (Manly et al., 1999, 2001) and is another common tool used to assess executive function in children. The children read a series of alternating numbers (e.g., 1, 2, 2, 1, 1, 2, 1, 2) aloud, in timed conditions. In the "same" condition, children read the numbers as they appear. In the "opposite" condition, children were asked to say the opposite of each digit (i.e., the previous sequence should be read as "2, 1, 1, 2, 2, 1, 2, 1"). An example is shown in **Figure 2**.

The variable of interest was the amount of time taken in the "opposite" condition, which requires inhibition of a prepotent verbal response: a faster response is taken to indicate an advantage in executive function.

Dimensional change card sort (DCCS)

A common measure of executive function in early childhood is the Dimensional Change Card Sort test (DCCS; Bialystok and Martin, 2004; see Zelazo, 2006 for the protocol adopted in this study). The standard version of this task requires children to sort a set of cards according to a particular dimension, such as color (e.g., "If it is blue it goes here, if it is red it goes there"); the children are subsequently asked to sort the same set of cards by according to a new dimension, such as shape (e.g., "If it is a rabbit it goes here, if it is a boat it goes there"). The test measures whether the child is able to switch from the first to the second dimension (marked as a 1), or instead, he/she keeps sorting the cards according to the first dimension (marked as 0). The variable of interest therefore is the number of correct responses.

Procedure

Written informed consent was obtained from parents of all participating children in accordance with the Declaration of Helsinki. The study was approved by the Linguistics and English Language ethics committee at the University of Edinburgh.

Testing took place during school hours in a quiet room made available by the schools. Each child was involved in two experimental sessions, with a gap of one day between sessions. In the first session, which lasted approximately 30 min, four tasks were administered to children the following order: COMPRENDO, Opposite Worlds, DCCS, and Raven. In the second session, which lasted approximately 15 min, children performed the remaining background tests: PPVT, Digit Span, and non-word repetition tasks. All children performed all the tests in the same order. All tests were administered in Italian to both bilingual and monolingual children.
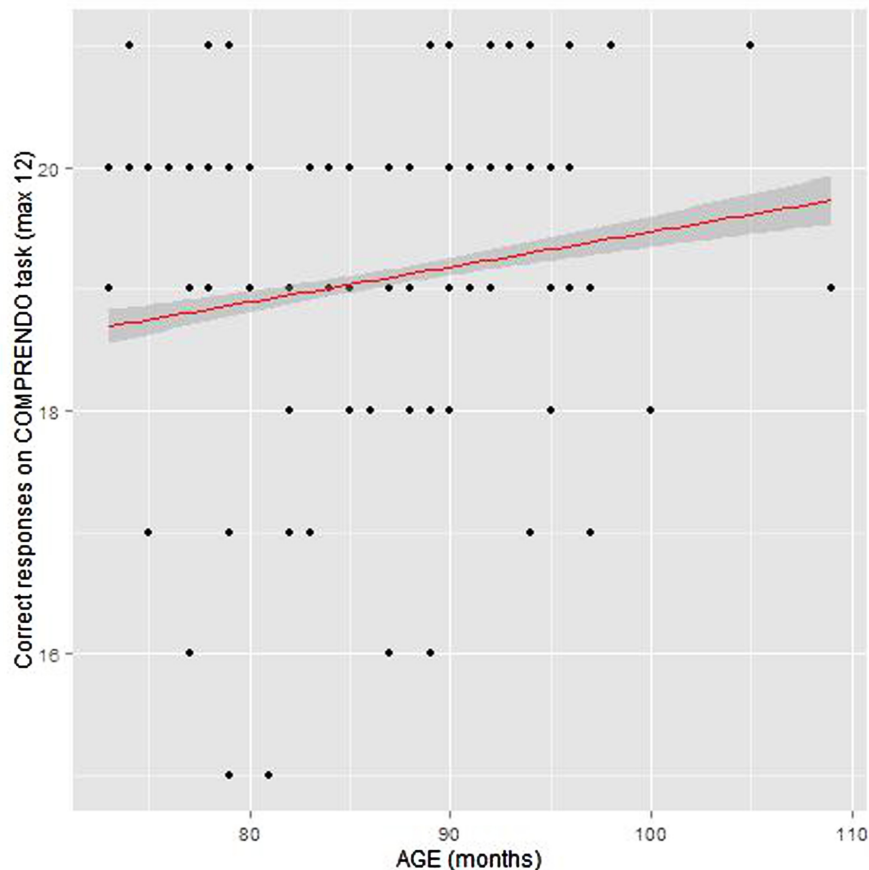
FIGURE 3 | Relationship between age and performance on COMPRENDO task.

Data Analyses
COMPRENDO

We used linear mixed effects (LME) models (e.g., Pinheiro and Bates, 2000) with logistic regression to estimate the likelihood of a correct response on a given trial. LME models with logistic regression have been demonstrated to handle categorical data (e.g., image selection) better than ANOVA (Jaeger, 2008). Mixed-effects modeling allows us to combine fixed effects (independent variables) with random effects terms sampled from a larger population, such as participant or item, thus capturing more of the random variance in a given data set (Baayen, 2008). All LME models were implemented in the lme4 package (Bates and Maechler, 2009) in R statistical software (R Development Core Team, 2011). All predictors were center prior to analysis, and coded using effects coding. This procedure helps to minimize collinearity (Baayen, 2008) and means that significance tests in the mixed-effects model correspond to tests for main effects and interactions in an ANOVA model (Cohen et al., 2003).

Opposite Worlds and DCCS Tasks

The opposite world task and DCCS produced a single statistic per child. Therefore, it was not possible to run LME models on these data, as random effects for participants or items were precluded. A standard linear model with age group, language group and their interactions as fixed effects was used instead.

RESULTS

Background Measures

A summary of mean ages, cumulative exposure to Italian, and scores on background measures (RAVEN, PPVT-4, Digit span test, and non-word repetition) for the four age groups of participants is given in **Table 2**.

TABLE 3 | Coefficients for linear mixed effects model in COMPRENDO: likelihood of correct response across all sentences combined ~ Age group * Language group.

	Estimate	SE	p
(Intercept)	2.88	0.30	<0.001***
Age group	0.44	0.19	<0.05*
Language group	−0.27	0.19	0.16
Age group: language group	−0.56	0.36	0.13

*P ≤ 0.05, ***P ≤ 0.001.

TABLE 4 | COMPRENDO: performance by sentence type and participant group.

	Active (%)	Passive (%)	Dative (%)	Coordinate (%)	Inflected (%)	Subject relative (%)	Object relative (%)
Bilingual Y1	100	96	94	87	85	85	76
Bilingual Y2	100	97	97	97	88	88	92
Monolingual Y1	97	100	95	93	75	92	73
Monolingual Y2	100	99	97	92	84	84	80
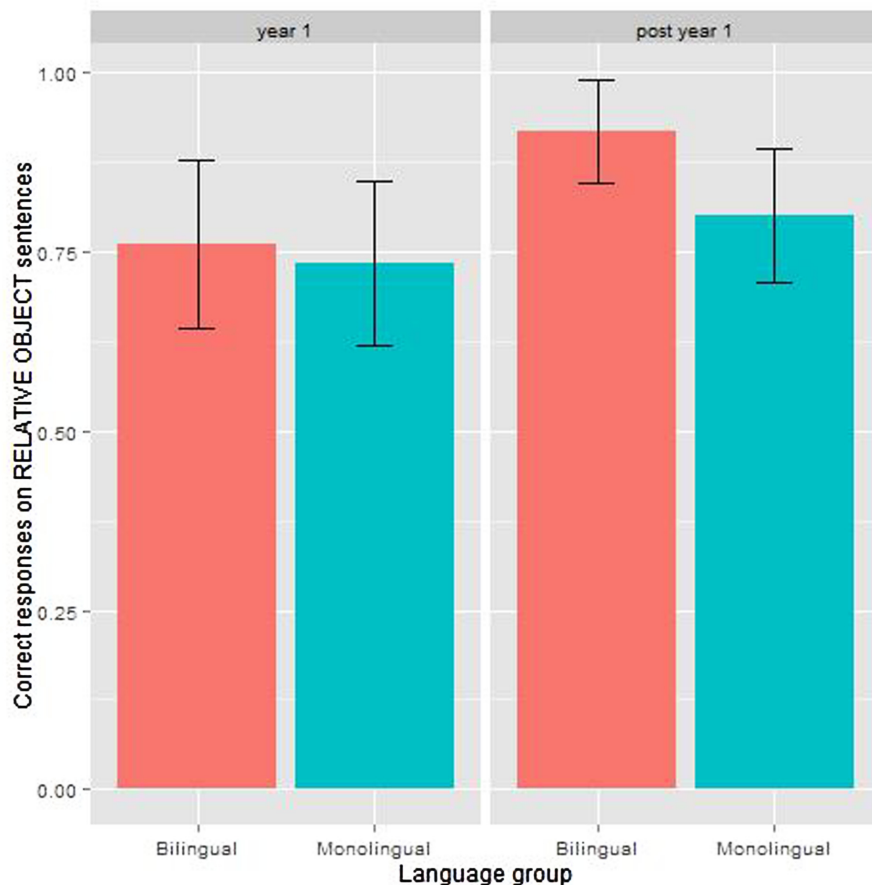
FIGURE 4 | Results for object relatives in the COMPRENDO test by language group and age group.

Linear models were used to test for significant differences between groups (language, age, and language by age) on the Raven CPM, PPVT-4, Digit span and non-word repetition background tests. Gaussian models were used for the Raven CPM, PPVT-4 and non-word repetition scores, and a Poisson model was used for the digit-span counts. Neither language (monolingual vs. bilingual), age (younger vs. older) or the interaction of language by age accounted for any significant difference in performance on the Raven CPM (Age-Language Group: est. -1.471, $SE = 1.585$, p. 0.36), PPVT-4 (Age-Language_Group: est. 4.78, $SE = 5.57$, p. 0.39, and Digit span tasks (Age-Language_Group: est. 0.12, $SE = 0.20$, p. 0.52). For the non-word repetition task, there was a main effect of age group, with the younger children making more errors than the older children (Age: est. 1.00, $SE = 0.5$, *$p < 0.05$), but no effect of language group or interaction between the two (Language_Group: est. 1.23, $SE = 0.68$, p. 0.07; Age-Language_Group: est -1.16, $SE = 1.01$, p. 0.25).

Test Measures
COMPRENDO
In the COMPRENDO task the children matched pictures to sentences of various levels of complexity. Recall that there were seven sentence types in total; active, passive, dative, coordinate, topicalised, subject relative, and object relative. We begin by analyzing all sentence types combined. We built an LME model using logistic regression. The dependent variable was the likelihood of a correct response on any given trial. The

TABLE 5 | Coefficients for linear mixed effects model in COMPRENDO: likelihood of correct response to object relative sentences ~ Age group * Language group.

	Estimate	SE	p
(Intercept)	1.66	0.45	<0.001***
Age group	0.84	0.36	<0.05*
Language group	−0.67	0.39	0.08
Age group: language group	−0.92	0.72	0.20

*P ≤ 0.05, ***P ≤ 0.001.

fixed effects were age group and language group, and their interactions. The model with maximal random effect structure failed to converge; this was a problem for all LME models in this section. We therefore removed the correlation parameter and the interaction term from the random slopes. This simplification resulted in a converged model and was used throughout these results unless otherwise specified.

The average correct responses (of a maximum 21) across all groups was 19.10 ($SD = 1.44$; 91% correct). The model showed that children in their first year of schooling were significantly more likely to give a correct response ($M = 18.79$, $SD = 1.54$; 89% correct) than those in their second year of schooling ($M = 19.34$, $SD = 1.31$; 92% correct) **Figure 3**.

Bilingual children scored higher ($M = 19.25$, $SD = 1.36$; 91%) than monolingual children ($M = 19.00$, $SD = 1.37$; 90%), but this

difference was not significant. The interaction between age group and language group was not significant. **Table 3** shows the model coefficients.

We then examined each type of sentence in turn. Performance by sentence type and by participant group is shown in **Table 4**. For active, passive, dative, coordinate, inflected, and subject relative sentences, there were no significant differences between age group or language group, and no significant interactions.

For object relative sentences, the model showed a significant effect of age: older children answered correctly on 85% trials, compared with 75% for younger children (see **Figure 4**). The bilingual group was more likely to give a correct response (84% correct answers) than the monolingual group (77% correct answers), however, this was only marginally significant. There was no significant interaction between age and language group. **Table 5** shows the model coefficients.

Opposite World Task

A linear model was built in which the dependent variable was the amount of time taken in the "opposite" condition. The fixed effects were age group (first year of schooling or second year of schooling), and language group (monolingual or bilingual), and their interaction.

The average time across all age and language groups was 41.9 s ($SD = 10.02$). As expected, speed on this task decreased with age: the older the child, the faster they performed the task (see
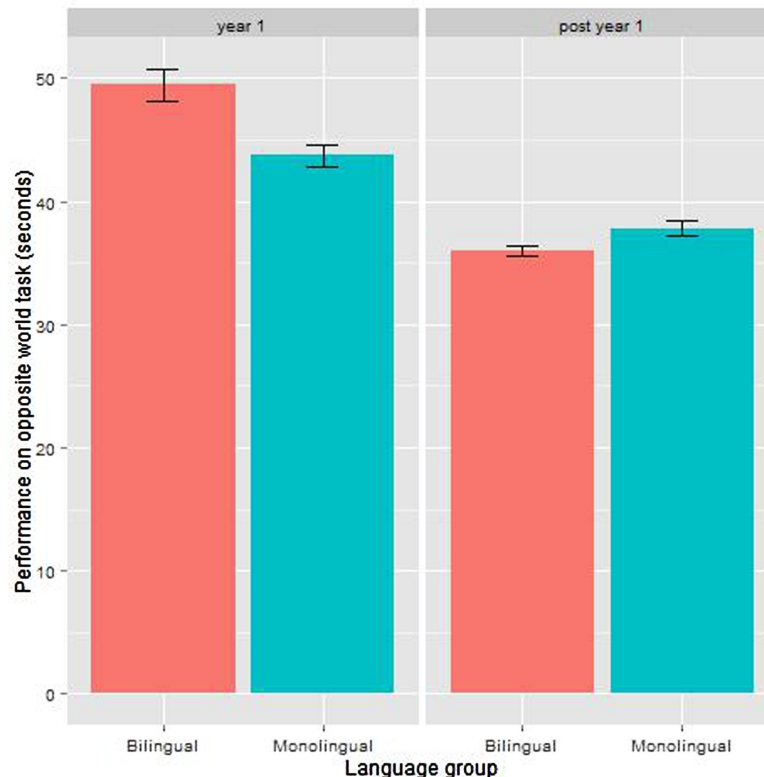
FIGURE 5 | Results in the Opposite World task by language group and age group.

TABLE 6 | Coefficients for linear model in opposite world: time taken in opposite world task ~ Age group * Language group.

	Estimate	SE	T	p
(Intercept)	41.18	0.20	202.03	<0.001***
Age group	−9.45	0.41	−23.05	<0.001***
Language group	−1.57	0.41	−3.85	<0.001***
Age group: Language group	7.54	0.82	9.19	<0.001***

***P ≤ 0.001.

Figure 5). The linear model showed a significant effect of age: the older age group performed faster ($M = 36.96$, $SD = 5.97$) than the younger age group ($M = 46.42$, $SD = 11.43$). There was also a significant effect of language group, with bilingual children being slightly slower ($M = 42.05$, $SD = 11.21$) than monolingual children ($M = 40.42$, $SE = 8.75$); and this is mainly due to the younger bilingual children's performance. The interaction between age group and language group is also significant: the bilingual children in their first year of schooling were 5.74 s slower on the task than their monolingual peers; bilingual children in their second year of schooling children were 1.8 s faster on the task than their monolingual peers. **Table 6**

shows the model coefficients. See **Table 8** for mean and SD by group.

DCCS Task

A linear model was built in which the dependent variable was the number of correct answers from a maximum of 12. The fixed effects were age group (first year of schooling or second year of schooling), and language group (monolingual or bilingual), and their interaction.

The average score across all groups was 8.57 ($SD = 2.3$). The linear model showed a significant effect of age, with children in their first year of schooling scoring lower ($M = 8.29$, $SD = 2.27$) than children in their second year of schooling ($M = 8.79$, $SD = 2.33$). There was also a significant effect of language group, with bilingual children scoring higher ($M = 9.03$, $SD = 2.23$) than monolingual children ($M = 8.16$, $SD = 2.32$); and this is mainly due to the older bilingual children's performance (see **Figure 6**). The interaction between age group and language group is significant: the monolingual children's score is more or less constant across years 1 and 2 of schooling, but the bilingual children in year 2 score higher than their bilingual peers in year 1. **Table 7** shows the model coefficients. See **Table 8** for mean and SD by group.
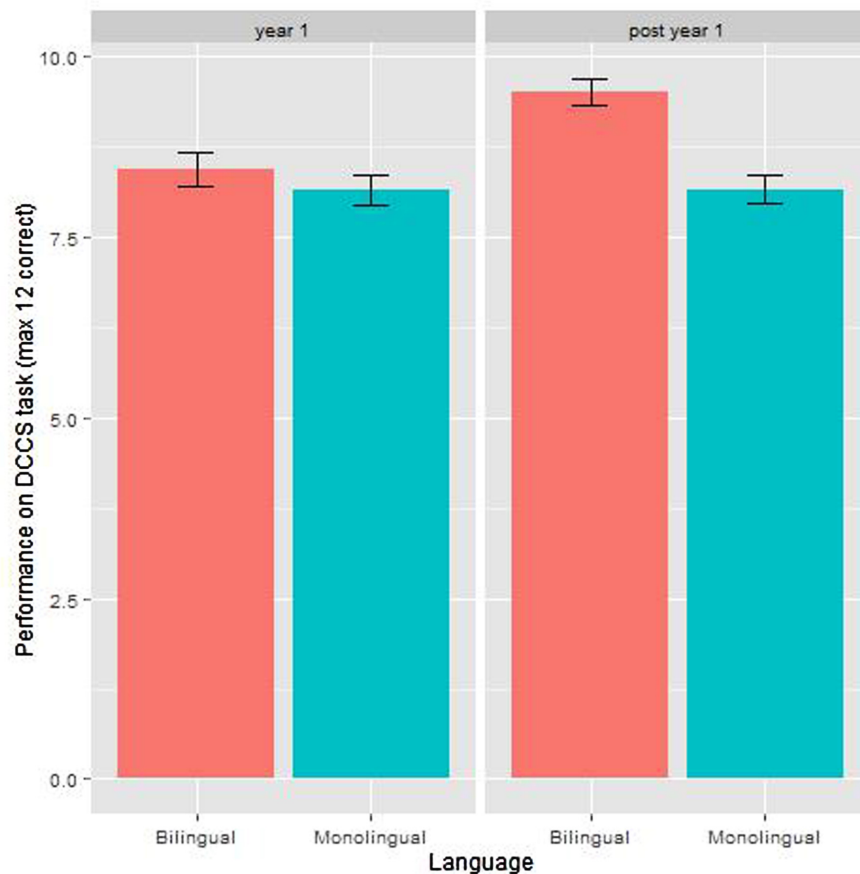
FIGURE 6 | Results of the DCCS by language group and age group.

DISCUSSION

The results of the study reported here can be summarized as follows:

(a) There are no significant differences between Sardinian–Italian bilingual children and monolingual Italian children in the control measures (i.e., Raven CPM, PPVT-4, Digit span, and non-word repetition).

(b) Overall, Sardinian–Italian bilingual children performed very similarly to monolingual Italian children in the COMPRENDO receptive grammatical test. All older children performed better than younger children, regardless of language group. There is a marginally significant difference in favor of bilinguals with respect to the comprehension of object relatives, which are the most complex of the seven syntactic structures tested: this difference is especially visible in older bilingual children.

(c) For the Opposite Worlds task, which is a test of executive functions requiring a verbal response, older children were overall faster than younger children. Younger bilingual children were slower than younger monolinguals whereas older bilinguals were faster than older monolinguals. This means that the score difference between younger and older children was wider for the bilingual group. Although the findings show an overall disadvantage for bilingual children, this is due to the large difference in performance between monolinguals and bilinguals in year 1, which is no longer present (and indeed, reversed – although not to the same extent) by year 2. The fact that this is a cross-sectional and not a longitudinal study invites caution in interpreting this difference as a steeper improvement in bilinguals. Furthermore, the verbal response required was in Italian, which may also have contributed to the monolingual advantage in younger children.

(d) For the DCCS, which is a task of executive function requiring a non-verbal response, there is improvement across the board from younger to older children. Younger bilinguals perform similarly to younger monolinguals. However, bilingual children provide more accurate responses in the older group.

These data reveal that bilingualism in Sardinian does not hinder development of linguistic competence in Italian, despite the fact that many of the bilingual children tested were dominant in Sardinian at the beginning of schooling. Bilingual children performed like monolinguals regardless of whether Sardinian and Italian are structurally similar or not. The trend toward bilingual advantages in comprehension of the object relative structure is more evident in older children. This may be regarded as further evidence that these advantages emerge gradually over time, as Bialystok et al. (2014) showed for children in immersion programs.

There is an alternative potential linguistic explanation for the trend toward a bilingual-monolingual difference in object relatives. In a study of adult learners of L2 Italian, Belletti and Guasti (2015) report that beginning L2 speakers are

TABLE 7 | Coefficients for linear model in DCCS: number of correct responses ~ Age group * Language group.

	Estimate	SE	t	p
(Intercept)	8.57	0.05	160.75	<0.001***
Age group	0.50	0.11	4.684	<0.001***
Language group	−0.87	0.11	−8.17	<0.001***
Age group : Language group	−1.05	0.21	−4.87	<0.001***

$***P \leq 0.001.$

TABLE 8 | Mean performance for opposite world and DCCS tasks by participant group (SD in parentheses).

	Opposite world (seconds)	DCCS (correct responses/12)
Bilingual Y1	49.44 (12.51)	8.44 (2.31)
Bilingual Y2	36.00 (4.40)	9.50 (2.04)
Monolingual Y1	43.70 (9.61)	8.15 (2.22)
Monolingual Y2	37.80 (6.97)	8.16 (2.40)

better than advanced L2 speakers, and often show ceiling performance in the production of object relatives. A very low percentage of passive object relatives were attested in beginner L2 speakers (22%) compared to a much higher production of passive object relatives in advanced L2 speakers (60%). In contrast, beginning L2 speakers produced 77% of correct object relatives compared to just 15% in the advanced group, approaching the performance of native Italian speakers. The low attested productions of passive object relatives in low proficiency Italian L2 speakers seems to mirror the finding of the present study that Sardinian–Italian bilingual children are marginally better at comprehending object relatives. Belletti and Guasti (2015) suggest that avoidance strategies are not available at early stages of acquisition in L2 speakers possibly due to a still imperfect command of morphosyntactic features of the language. It is unclear how avoidance strategies would affect comprehension. Notice, however, that 'imperfect command' here is not necessarily to be understood as lack of relevant knowledge, but possibly as slower access to alternative structures that may compete with object relatives. It is also possible that bilingual children may have sufficient inhibitory control to exclude the alternative structures. These differences cannot be directly tested in this study, and further research is necessary to explore these alternative accounts.

The analysis of the cognitive test results points to a global improvement from younger to older children, and to an overall advantage for older bilingual children. The Opposite World test and the DCCS both test aspects of executive functions, such as the ability to inhibit an inappropriate response and switch between conditions. Only the Opposite World test, however, requires an overt verbal response (in Italian). It is in this test that younger bilingual children (whose home language is Sardinian, rather than Italian) have an initial disadvantage compared to monolinguals. In the DCCS, on the other hand, bilinguals and monolinguals are the same in the younger group. A plausible

interpretation of this disparity between the tests may be related to the fact that the bilinguals in the first year of primary school had experienced comparably fewer opportunities to use Italian productively. As was the case for the COMPRENDO test, advantages in cognitive function may emerge gradually with time and more exposure to both languages. In any case, bilingualism involving a regional minority language may come with some of the same beneficial effects as bilingualism in other languages.

CONCLUSION

This study involved 85 children from the Nuoro province of central Sardinia, of whom 45 were monolingual in Italian and 40 were bilingual in Sardinian and Italian. All children were comparable with respect to vocabulary knowledge, phonological memory, typical language development, and general intelligence. The children performed in a test of Italian receptive competence and in two standardized tests of executive functions. In most cases the performance of bilingual children was not different from monolinguals.

This study has limitations. The most obvious ones are the limited size of the sample, the cross-sectional design, and the narrow range of abilities tested. Future research will explore the relationship between comprehension and production abilities in the Italian of Sardinian–Italian children, as well as the correlations between language abilities and cognitive abilities. The full range of abilities should be studied over a longer period of time, in both longitudinal and cross-sectional studies, to establish the developmental trajectories of both linguistic and general cognitive skills, and of the effects on each other. Despite these limitations, however, the results of this study are inconsistent with the common perception that bilingualism with Sardinian is a cognitive burden and compromises performance in Italian.

REFERENCES

Adani, F. (2011). Re-thinking the acquisition of relative clauses in Italian: towards a grammatically-based account. *J. Child Lang.* 22, 1–25. doi: 10.1017/S0305000909990250

Adani, F., van der Lely, H. K., Forgiarini, M., and Guasti, M. T. (2010). Grammatical feature dissimilarities make relative clauses easier: a comprehension study with Italian children. *Lingua* 120, 2148–2166. doi: 10.1016/j.lingua.2010.03.018

Antoniou, K., Kambanaros, M., Grohmann, K., and Katsos, N. (2014). Is bilectalism similar to bilingualism? An investigation into children's vocabulary and executive control skills. *Proc. BUCLD* 38, 12–24.

Baayen, R. H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics.* Cambridge: Cambridge University Press.

Bates, D., and Maechler, M. (2009). *Lme4: Linear Mixed-Effects Models Using S4 Classes. R Package Version 0.999375–999331.*

Baum, S., and Titone, D. (2014). Moving towards a neuroplasticity view of bilingualism, executive control, and aging. *Appl. Psychol.* 35, 857–894. doi: 10.1017/S0142716414000174

Belletti, A., and Guasti, M. T. (2015). *The Acquisition of Italian. Morphosyntax and its Interfaces in Different Modes of Acquisition.* Amsterdam: John Benjamins.

Bernardini, P., and Schlyter, S. (2004). Growing syntactic structure and code-mixing in the weaker language: the Ivy Hypothesis. *Biling. Lang. Cogn.* 7, 49–69. doi: 10.1017/S1366728904001270

Bialystok, E. (1988). Levels of bilingualism and levels of linguistic awareness. *Dev. Psychol.* 24, 560–567. doi: 10.1037/0012-1649.24.4.560

Bialystok, E. (1992). Attentional control in children's metalinguistic performance and measures of field independence. *Dev. Psychol.* 28, 654–664. doi: 10.1037/0012-1649.28.4.654

Bialystok, E. (2009). Bilingualism: the good, the bad, and the indifferent. *Biling. Lang. Cogn.* 12, 3–11. doi: 10.1017/S1366728908003477

Bialystok, E., Craik, F. I., and Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends Cogn. Sci.* 16, 240–250. doi: 10.1016/j.tics.2012.03.001

Bialystok, E., Luk, G., Peets, K. F., and Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Biling. Lang. Cogn.* 13, 525–531. doi: 10.1017/S1366728909990423

Bialystok, E., and Martin, M. (2004). Attention and inhibition in bilingual children: evidence from the dimensional change card sort task. *Dev. Psychol.* 7, 325–339.

Bialystok, E., Peets, K., and Moreno, S. (2014). Producing bilinguals through immersion education: development of metalinguistic awareness. *Appl. Psychol.* 35, 177–191. doi: 10.1017/S0142716412000288

Blom, E. (2010). Effects of input on the early grammatical development of bilingual children. *Int. J. Biling.* 14, 422–446. doi: 10.1177/1367006910370917

Blumenfeld, H., and Marian, V. (2011). Bilingualism influences inhibitory control in auditory comprehension. *Cognition* 118, 245–257. doi: 10.1016/j.cognition.2010.10.012

Bolognesi, R. (2013). *Le Identita' Linguistiche dei Sardi.* Cagliari: Condaghes.

Cecchetto, C., Di Domenico, A., Garraffa, M., and Papagno, C. (2012). *Comprendo: Batteria per la Comprensione di Frasi Negli Adulti*. Milan: Raffaello Cortina Editore.

Cohen, J., Cohen, P., West, S. G., and Aiken, L. (2003). *Applied Multiple Regression/Correlation Analysis in the Behavioral Sciences*, 3rd Edn. Mahwah, NJ: Erlbaum.

Collins, C. (2005). A smuggling approach to the passives in English. *Syntax* 82, 81–120. doi: 10.1111/j.1467-9612.2005.00076.x

Contemori, C., and Belletti, A. (2012). Relatives and passive object relatives in Italian-speaking children and adults: intervention in production and comprehension. *Appl. Psychol.* 35, 1021–1053. doi: 10.1017/S0142716412000689

Contemori, C., and Garraffa, M. (2010). Comparison of modalities in SLI syntax: a study on the comprehension and production of non-canonical sentences. *Lingua* 120, 1940–1955. doi: 10.1016/j.lingua.2010.02.011

Cornoldi, C., Miato, L., Molin, A., and Poli, S. (2009). *PRCR-2. Prove di Prerequisito per la Diagnosi Delle Difficoltà di Lettura e Scrittura*. Firenze: Giunti.

Costa, A., Hernández, M., Costa-Faidella, J., and Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: now you see it, now you don't. *Cognition* 113, 135–149. doi: 10.1016/j.cognition.2009.08.001

Costa, A., Santesteban, M., and Ivanova, I. (2006). How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 1057–1074.

Costa, A., and Sebastian-Galles, N. (2014). How does the bilingual experience sculpt the brain? *Nat. Rev. Neurosci.* 15, 336–345. doi: 10.1038/nrn 3709

Craik, F., and Bialystok, E. (2006). Cognition through the lifespan: mechanisms of change. *Trends Cogn. Sci.* 10, 131–138. doi: 10.1016/j.tics.2006.01.007

Crystal, D. (2000). *Language Death*. Cambridge: Cambridge University Press.

de Bruin, A., Treccani, B., and Della Sala, S. (2015). Cognitive advantage in bilingualism: an example of publication bias? *Psychol. Sci.* 26, 99–107. doi: 10.1177/0956797614557866

De Houwer, A. (1990). *The Acquisition of Two Languages from Birth: A Case Study*. Cambridge: Cambridge University Press.

Dopke, S. (1998). Competing language structures: the acquisition of verb placement by bilingual German-English children. *J. Child Lang.* 25, 555–584. doi: 10.1017/S0305000998003584

Duñabeitia, J. A., Hernández, J. A., Antón, E., Macizo, P., Estévez, A., Fuentes, L. J., et al. (2014). The inhibitory advantage in bilingual children revisited: myth or reality? *Exp. Psychol.* 61, 234–251. doi: 10.1027/1618-3169/a000243

Extra, G., and Gorter, D. (2008). *Multilingual Europe: Facts and Policies*. Berlin: Mouton De Gruyter.

Friedmann, N., Belletti, A., and Rizzi, L. (2009). Relativized relatives: types of intervention in the acquisition of A-bar dependencies. *Lingua* 119, 67–88. doi: 10.1016/j.lingua.2008.09.002

Garraffa, M., and Grillo, N. (2008). Canonicity effects as grammatical phenomena. *J. Neurol.* 21, 177–197. doi: 10.1016/j.jneuroling.2007.09.001

Gathercole, S. (1998). The development of memory. *J. Child Psychol. Psychiatry* 39, 3–27. doi: 10.1111/j.1467-7687.2009.00908.x

Gathercole, V., Thomas, E., Kennedy, I., Prys, C., Young, N., Viñas Guasch, N., et al. (2014). Does language dominance affect cognitive performance in bilinguals? Lifespan evidence from preschoolers through older adults on card-sorting, Simon, and metalinguistic tasks. *Front. Psychol.* 5:11. doi: 10.3389/fpsyg.2014.00011

Gathercole, V. C. M. (2007). Miami and north Wales, so far and yet so near: constructivist account of morpho-syntactic development in bilingual children. *Int. J. Biling. Educ. Biling.* 10, 224–246. doi: 10.2167/beb442.0

Genesee, F., Nicoladis, E., and Paradis, J. (1995). Language differentiation in early bilingual development. *J. Child Lang.* 22, 611–631. doi: 10.1017/S0305000900009971

Goetz, P. (2003). The effects of bilingualism on theory of mind development. *Biling. Lang. Cogn.* 6, 1–15. doi: 10.1016/j.tics.2008.02.001

Gordon, A. C. L., and Olson, D. R. (1998). The relation between acquisition of a theory of mind and the capacity to hold in mind. *J. Exp. Child Psychol.* 68, 70–83. doi: 10.1006/jecp.1997.2423

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Biling. Lang. Cogn.* 1, 67–81. doi: 10.1017/S1366728998000133

Green, D. W., and Abutalebi, J. (2013). Language control in bilinguals: the adaptive control hypothesis. *J. Cogn. Psychol.* 25, 515–530. doi: 10.1080/20445911.2013.796377

Grenoble, L. A., and Whaley, L. J. (2006). *Saving Languages: An Introduction to Language Revitalization*. Cambridge: Cambridge University Press.

Grillo, N., (2008). *Generalized Minimality*. Ph.D. thesis, Utrecht Institute of Linguistics OTS, Utrecht.

Grohmann, K. K. (2014). Towards comparative bilingualism. *Ling. Approach. Biling.* 4, 336–341. doi: 10.1075/lab.4.3.06gro

Guasti, M. T., Maggioni, A., and Vernice M. (2013). "Disturbi specifici del linguaggio, bilinguismo e acquisizione di L2," in *Nuovi Contesti d'acquisizione e Insegnamento: l'italiano Nelle Realtà Plurilingui. Atti del Convegno CIS 2012 – Bergamo*, Vol. 9, ed. A. Grassi (Perugia: Guerra Edizioni).

Harris, M., and Vincent, N. (1988). *The Romance Languages*. London: Routledge.

Hilchey, M. D., and Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychon. Bull. Rev.* 4, 625–658. doi: 10.3758/s13423-011-0116-7

Houston, D. M., and Jusczyk, P. (2000). The role of talker-specific information in word segmentation by infants. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 1570–1582.

Jaeger, T. F. (2008). Categorical data analysis: away from ANOVAs (transformation or not) and towards logit mixed models. *J. Mem. Lang.* 59, 434–446. doi: 10.1016/j.jml.2007.11.007

Jones, M. A. (1993). *Sardinian Syntax*. London: Psychology Press.

Kovács, A., and Mehler, J. (2009). Cognitive gains in 7-month old bilingual infants. *Proc. Natl. Acad. Sci. U.S.A.* 106, 6556–6560. doi: 10.1073/pnas.0811323106

Lauchlan, F., Parisi, M., and Fadda, R. (2013). Bilingualism in sardinia and scotland: exploring the cognitive benefits of speaking a 'minority' language. *Int. J. Biling.* 17, 43–56. doi: 10.1177/1367006911429622

Manly, T., Anderson, V., Nimmo-Smith, I., Turner, A., Watson, P., and Robertson, I. H. (2001). The differential assessment of children's attention: the test of everyday attention for children (TEA-Ch), normative sample and ADHD performance. *J. Child Psychol. Psychiatry* 42, 1065–1081. doi: 10.1111/1469-7610.00806

Manly, T., Robertson, I. H., Anderson, V., and Nimmo-Smith, I. (1999). *The Test of Everyday Attention for Children*. Bury St. Edmunds: Thames Valley Test Company Limited.

Meisel, J. (2007). The weaker language in early child bilingualism: acquiring a first language as a second language? *Appl. Psychol.* 28, 495–514. doi: 10.1017/S0142716407070270

Meisel, J. M. (1989). "Early differentiation of languages in bilingual children," in *Bilingualism Across the Lifespan: Aspects of Acquisition, Maturity and Loss*, eds K. Hyltenstam and L. Obler (Cambridge: Cambridge University Press), 13–40.

Miyake, A., and Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: four general conclusions. *Curr. Dir. Psychol. Sci.* 21, 8–14. doi: 10.1177/0963721411429458

Müller, N., and Hulk, A. (2001). Crosslinguistic influence in bilingual language acquisition: Italian and french as recipient languages. *Biling. Lang. Cogn.* 4, 1–21.

Nettle, D. (1999). *Linguistic Diversity*. Oxford: Oxford University Press.

Oppo, A. (2007). *Le Lingue dei Sardi: Una Ricerca Sociolinguistica*. Cagliari: Regione Autonoma della Sardegna.

Orsini, A., Grossi, D., Capitani, E., Laiacona, M., Papagno, C., and Vallar, G. (1987). Verbal and spatial immediate memory span: normative data from 1355 adults and 1112 children. *Ital. J. Neurol. Sci.* 8, 537–548. doi: 10.1007/BF023 33660

Paap, K. R. (2014). The role of componential analysis, categorical hypothesizing, replicability and confirmation bias in testing for bilingual advantages in executive functioning. *J. Cogn. Psychol.* 26, 242–255. doi: 10.1080/20445911.2014.891597

Paap, K. R., and Greenberg, Z. L. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cogn. Psychol.* 66, 232–258. doi: 10.1016/j.cogpsych.2012.12.002

Paradis, J., and Navarro, S. (2003). Subject realization and crosslinguistic interference in the bilingual acquisition of spanish and english: what is the role of the input? *J. Child Lang.* 30, 371–393. doi: 10.1017/S0305000903005609

Paradis, J., Nicoladis, E., Crago, M., and Genesee, F. (2011). Bilingual children's acquisition of the past tense: a usage-based approach. *J. Child Lang.* 38, 554–578. doi: 10.1017/S0305000910000218

Pinheiro, J. C., and Bates, D. (2000). *Mixed Effects Models in S and S-plus*. Berlin: Springer.

Posner, R. (1996). *The Romance Languages*. Cambridge: Cambridge University Press.

Prior, A., and Gollan, T. H. (2011). Good language-switchers are good task-switchers: evidence from Spanish-English and Mandarin-English bilinguals. *J. Int. Neuropsychol. Soc.* 17, 682–691. doi: 10.1017/S1355617711000580

R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available at: http://www.r-project.org [accessed January 31, 2011].

Raven, J., Court, J. H., and Raven, J. C. (1998). *Raven Manual, Section 1 (General overview) and Section 2 (Coloured Progressive Matrices)*. Oxford: Oxford Psychologist Press.

Romaine, S. (2007). Preserving endangered language. *Lang. Ling. Compass* 1, 115–132. doi: 10.1111/j.1749-818X.2007.00004.x

Serratrice, L. (2007). Cross-linguistic influence in the interpretation of anaphoric and cataphoric pronouns in English-Italian bilingual children. *Biling. Lang. Cogn.* 10, 225–238. doi: 10.1017/S1366728907003045

Serratrice, L., Sorace, A., and Paoli, S. (2004). Crosslinguistic influence at the syntax-pragmatics interface: subjects and objects in Italian-English bilingual and monolingual acquisition. *Biling. Lang. Cogn.* 7, 183–206. doi: 10.1017/S1366728904001610

Siegal, M., Iozzi, L., and Surian, L. (2009). Bilingualism and conversational understanding in young children. *Cognition* 110, 115–122. doi: 10.1016/j.cognition.2008.11.002

Siegal, M., Surian, L., Matsuo, A., Geraci, A., Iozzi, L., Okumura, Y., et al. (2010). Bilingualism accentuates children's conversational understanding. *PLoS ONE* 5:e9004. doi: 10.1371/journal.pone.0009004

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. *Ling. Approach. Biling.* 1, 1–33. doi: 10.1075/lab.1.1.01sor

Sorace, A. (2012). Pinning down the concept of "interface" in bilingualism: a reply to peer commentaries. *Ling. Approach. Biling.* 2, 209–216. doi: 10.1075/lab.2.2.04sor

Sorace, A., and Serratrice, L. (2009). Internal and external interfaces in bilingual language development: beyond structural overlap. *Int. J. Biling.* 13, 195–210. doi: 10.1177/1367006909339810

Stella, G., Pizzoli, C., and Tressoldi, P. (2000). *Peabody. Test Psicolinguistico*. Torino: Omega Edizioni.

Tamburelli, M., Sanoudaki, E., Jones, G., and Sowinska, M. (2015). Acceleration in the bilingual acquisition of phonological structure: evidence from Polish–English bilingual children. *Biling. Lang. Cogn.* 18, 713–725. doi: 10.1017/S1366728914000716

Tucker, G. R. (1998). "A global perspective on multilingualism and multilingual education," in *Beyond bilingualism: Multilingualism and Multilingual Education*, eds J. Cenoz and F. Genesee (Clevedon: Multilingual Matters), 3–15.

Unsworth, S. (2005). *Child L2, Adult L2, Child L1: Differences and Similarities. A Study on the Acquisition of Direct Object Scrambling in Dutch*. PhD thesis, Utrecht University, Utrecht.

Unsworth, S. (2013a). Assessing age of onset effects in (early) child L2 acquisition. *Lang. Acquis.* 20, 74–92. doi: 10.1080/10489223.2013.766739

Unsworth, S. (2013b). Current issues in multilingual first language acquisition. *Annu. Rev. Appl. Ling.* 33, 21–50. doi: 10.1017/S0267190513000044

Unsworth, S., Argyri, F., Cornips, L., Hulk, A. C. J., Sorace, A., and Tsimpli, I. (2014). On the role of age of onset and input in early child bilingualism in Greek and Dutch. *Appl. Psychol.* 35, 765–805. doi: 10.1017/S0142716412000574

Valian, V. (2015). Bilingualism and cognition. *Biling. Lang. Cogn.* 18, 3–24. doi: 10.1017/S1366728914000698

Vangsnes, Ø., Söderlund, G. B. W., and Blekesaune, M. (2015). The effect of bidialectal literacy on school achievement. *Int. J. Biling. Educ. Biling.* 1–16. doi: 10.1080/13670050.2015.1051507

Yip, V., and Matthews, S. (2007). *The Bilingual Child: Early Development and Language Contact*. Cambridge: Cambridge University Press.

Zelazo, P. D. (2006). The dimensional change card sort (DCCS): a method of assessing executive function in children. *Nat. Protoc.* 1, 297–301. doi: 10.1038/nprot.2006.46

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read, for greatest visibility

COLLABORATIVE PEER-REVIEW
Designed to be rigorous – yet also collaborative, fair and constructive

FAST PUBLICATION
Average 85 days from submission to publication (across all journals)

COPYRIGHT TO AUTHORS
No limit to article distribution and re-use

TRANSPARENT
Editors and reviewers acknowledged by name on published articles

SUPPORT
By our Swiss-based editorial team

IMPACT METRICS
Advanced metrics track your article's impact

GLOBAL SPREAD
5'100'000+ monthly article views and downloads

LOOP RESEARCH NETWORK
Our network increases readership for your article

Frontiers
EPFL Innovation Park, Building I • 1015 Lausanne • Switzerland
Tel +41 21 510 17 00 • Fax +41 21 510 17 01 • info@frontiersin.org
www.frontiersin.org

Find us on