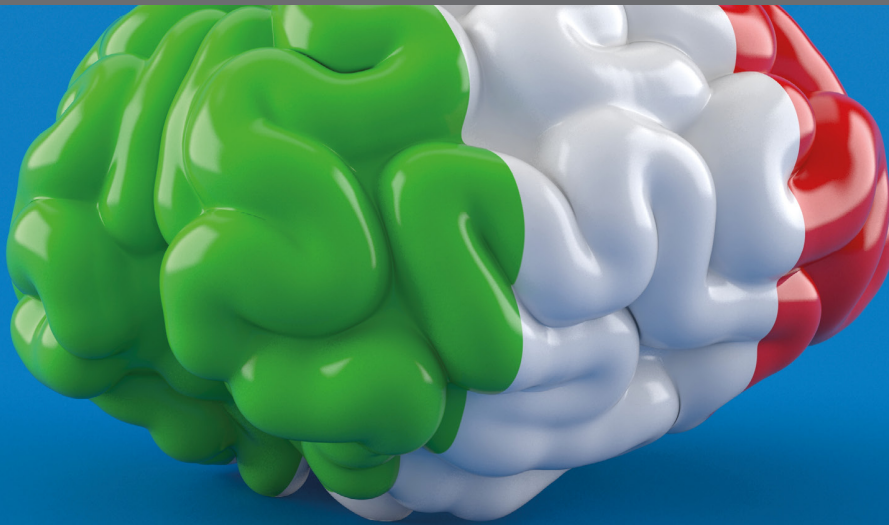# (PUSHING) THE LIMITS OF NEUROPLASTICITY INDUCED BY ADULT LANGUAGE ACQUISITION

EDITED BY: Jurriaan Witteman, Leticia Pablos-Robles, Maria Carmen Parafita Couto, Niels Schiller, Yiya Chen and Patrick Wong

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

(PUSHING) THE LIMITS OF NEUROPLASTICITY INDUCED BY ADULT LANGUAGE ACQUISITION

Topic Editors:
Jurriaan Witteman, Leiden University, Netherlands
Leticia Pablos-Robles, Leiden University, Netherlands
Maria Carmen Parafita Couto, Leiden University, Netherlands
Niels Schiller, Leiden University, Netherlands
Yiya Chen, Leiden University, Netherlands
Patrick Wong, The Chinese University of Hong Kong, Hong Kong

Image by: Talaj/Shutterstock.com

Most adults attempt to learn a second or even third language at some point in their life. Since language exposure is one of the most intense cognitive training regimes one can encounter, it is not surprising that previous research has shown that multilingualism can induce profound change in the brain or 'neuroplasticity'. What remains unclear is the scope of such adult language learning induced neuroplasticity. In other words, much is yet to be investigated about the factors that limit or promote adult language learning induced neuroplasticity.

On the one hand, the present research topic discusses research that sheds light on neural mechanisms that limit adult language learning induced neuroplasticity such as: neural mechanisms of first language interference in the acquisition of a second language and reduced opportunity for language induced neuroplasticity due to aging.

On the other hand, the Research Topic discusses factors that could enhance non-native language learning (and underlying neuroplastic mechanisms), such as the duration of the training regime, language aptitude, and meta-linguistic awareness.

Therefore, the goal of the present Research Topic is to examine both the limits of neuroplasticity in adult language learning and the ways to push beyond those limits. Understanding of such limits and frontiers to push beyond the limits is not only theoretically fundamental but could also have practical implications for enhancing language training programmes.

Table of Contents

Editorial: (Pushing) the Limits of Neuroplasticity Induced by Adult Language Acquisition

Jurriaan Witteman [1,2,3*], Yiya Chen [1,2,3], Leticia Pablos-Robles [1,2,3],
Maria Carmen Parafita Couto [1,2,3], Patrick C. M. Wong [4] and Niels O. Schiller [1,2,3]

[1] Department of Linguistics, Leiden University, Leiden, Netherlands, [2] Leiden University Centre for Linguistics, Leiden, Netherlands, [3] Leiden Institute for Brain and Cognition, Leiden University, Leiden, Netherlands, [4] Department of Linguistics and Modern Languages, Brain and Mind Institute, The Chinese University of Hong Kong, Hong Kong, China

Editorial on the Research Topic

(Pushing) the Limits of Neuroplasticity Induced by Adult Language Acquisition

Many individuals attempt to learn a second (L2) or even third language (L3) at some point in their life. Since language exposure is one of the most intense cognitive training regimes one can encounter, it is not surprising that previous research has shown that multilingualism can induce profound neural changes or "neuroplasticity" (Costa and Sebastián-Gallés, 2014). Despite the general consensus that learning a new language in adulthood can change the brain, what remains unclear is the *scope* of such neuroplasticity. In other words, what limits vs. promotes neurocognitive change as a result of second language acquisition in adulthood?

On the one hand, there are factors that may limit such change of the neurocognitive system due to L2 (or L3) acquisition. For instance, models of adult L2 learning assume that acquisition of the mother tongue (L1) has sculpted neural circuits to discriminate between L1 linguistic elements which in turn limits the ability to distinguish between L2 elements (e.g., van Leussen and Escudero, 2015). On the other hand, there might be factors that enhance L2 induced neurocognitive change, such as language aptitude (Hu et al., 2013; Chai et al., 2016) and the intensity (Tremblay et al., 1997; Thomson and Derwing, 2015) and quality (Zhang et al., 2009; Ylinen et al., 2010; Morgan-Short et al., 2012; Grimaldi et al., 2014) of the L2 acquisition regime. Hence, much is yet to be investigated about the factors that limit vs. promote adult language learning induced neuroplasticity as well as the mediating underlying neurocognitive mechanisms. The present research topic therefore aimed to identify some of the factors that limit or promote adult L2 learning induced neurocognitive plasticity and the underlying neurocognitive mechanisms.

What factors then, might limit neurocognitive change due to adult L2 acquisition? The two reviews presented in the current research topic (Birdsong; Antoniou and Wright) both suggest that having reached adult age itself might be a limiting factor because adult age represents a period of relatively (as compared to childhood) low susceptibility to L2 exposure, limiting the degree to which L2 proficiency can be gained. Additionally, the mismatch of L1-L2 typology was suggested to limit L2 acquisition, with a relatively large mismatch delaying successful L2 acquisition (Antoniou and Wright). Indeed, a cross-linguistic priming study with concurrent ERP recordings presented in the current research topic showed that already early in the L2 acquisition process there is interaction between L1 and L2 at the semantic level (Meade et al.). Another study (Yang et al.), showed that a large typological difference between L1 and an L3 makes switching between languages in bilinguals more difficult with different underlying cognitive control networks being engaged in switching between balanced vs. unbalanced languages.

Acquisition of L2 (and underlying neurocognitive change) may additionally vary *between* domains of the language or even *within* a domain. Indeed, a study presented in the current research topic showed that while learning an artificial language, words that correspond to relatively concrete concepts are more easily integrated into existing semantic networks than words that refer to relatively abstract concepts (Ding et al.).

Finally, as outlined above (adult) age may itself limit neuroplastic change due to L2 learning. Indeed, a structural white matter imaging study presented in the current research topic suggests that white matter bundles critical for obtaining syntactic abilities are still developing in adolescence but may have reached maturation in adults (Yamamoto and Sakai), perhaps limiting acquisition of L2 syntax at adult age. On the other hand, the limiting effect of age of acquisition on L2 induced structural neuroplasticity may in itself be limited, as shown by a study presented in the current research topic, demonstrating white matter differences purportedly due to L2 learning between mono- and bilinguals, despite of L2 learners having reached adulthood (Rossi et al.).

Having discussed the factors that may limit neuroplasticity due to adult L2 acquisition, what factors may promote it? One of the reviews of the present research topic (Birdsong) mentions some possible factors such as: high working memory capacity, motivation to learn and meta-linguistic awareness (which could be promoted by having successfully acquired a previous non-native language). Indeed, a study in the present research topic examining predictors of L2 acquisition success found evidence that high working memory capacity predicts L2 acquisition success (Blumenfeld et al.). Furthermore, the general ability

to learn or "language aptitude" may enhance neurocognitive change induced by L2 or L3 learning. Indeed, a study presented in our research topic investigating the morphology of Heschl's gyrus (HG), the primary auditory cortex, suggests that the number of complete duplications of HG in the right hemisphere might be a structural correlate of language aptitude, that may enhance L2 acquisition success (Turker et al.). Finally, in an interesting study examining the effects of L2 acquisition on *L1* processing presented in the current research topic (Kasparian and Steinhauer), very extended exposure to L2 and resulting high L2 proficiency emerged as an important factor in determining (abnormal) morphosyntactic L1 processing, suggesting that the intensity of L2 exposure is a critical determinant of neuroplastic change in the underlying neurocognitive architecture of the language processing system.

In sum, the studies presented in the current research topic suggest that neuroplastic change due to acquisition of another language (L2, L3, etc.) seems to be limited by adult age, typological mismatch between the already acquired and to be acquired languages, and limited exposure to the to be acquired language. On the other hand, high working memory capacity, high "language aptitude," and a high level of exposure to the to be acquired language seem to promote neuroplastic change. Together, we have aimed with the studies presented in the current research topic to provide a fresh look at the scope of neuroplastic change due to adult second language acquisition.

AUTHOR CONTRIBUTIONS

Conceived the idea: JW and NS; Wrote the Research Topic proposal: JW, YC, LP-R, MP, PW,and NS; Edited articles: JW, YC, LP-R, MP, PW, and NS; Wrote the editorial: JW.

REFERENCES

Chai, X. J., Berken, J. A., Barbeau, E. B., Soles, J., Callahan, M., Chen, J.-K., et al. (2016). Intrinsic functional connectivity in the adult brain and success in second-language learning. *J. Neurosci.* 36, 755–761. doi: 10.1523/JNEUROSCI.2234-15.2016

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

Grimaldi, M., Sisinni, B., Gili Fivela, B., Invitto, S., Resta, D., Alku, P., et al. (2014). Assimilation of L2 vowels to L1 phonemes governs L2 learning in adulthood: a behavioral and ERP study. *Front. Hum. Neurosci.* 8:279. doi: 10.3389/fnhum.2014.00279

Hu, X., Ackermann, H., Martin, J. A., Erb, M., Winkler, S., and Reiterer, S. M. (2013). Language aptitude for pronunciation in advanced second language (L2) Learners: Behavioural predictors and neural substrates. *Brain Language* 127, 366–376. doi: 10.1016/j.bandl.2012.11.006

Morgan-Short, K., Steinhauer, K., Sanz, C., and Ullman, M. T. (2012). Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *J. Cogn. Neurosci.* 24, 933–947. doi: 10.1162/jocn_a_00119

Thomson, R. I., and Derwing, T. M. (2015). The effectiveness of L2 pronunciation instruction: a narrative review. *Appl. Linguist.* 36, 326–344. doi: 10.1093/applin/amu076

Tremblay, K., Kraus, N., Carrell, T. D., and McGee, T. (1997). Central auditory system plasticity: generalization to novel stimuli following listening training. *J. Acoust. Soc. Am.* 102, 3762–3773.

van Leussen, J.-W., and Escudero, P. (2015). Learning to perceive and recognize a second language: the L2LP model revised. *Front. Psychol.* 6:1000. doi: 10.3389/fpsyg.2015.01000

Ylinen, S., Uther, M., Latvala, A., Vepsäläinen, S., Iverson, P., Akahane-Yamada, R., et al. (2010). Training the brain to weight speech cues differently: a study of finnish second-language users of english. *J. Cogn. Neurosci.* 22, 1319–1332. doi: 10.1162/jocn.2009.21272

Zhang, Y., Kuhl, P. K., Imada, T., Iverson, P., Pruitt, J., Stevens, E. B., et al. Nemoto, I. (2009). Neural signatures of phonetic learning in adulthood: A magnetoencephalography study. *Neuroimage* 46, 226–240. doi: 10.1016/j.neuroimage.2009.01.028

Plasticity, Variability and Age in Second Language Acquisition and Bilingualism

David Birdsong*

Department of French and Italian, The University of Texas at Austin, Austin, TX, United States

Much of what is known about the outcome of second language acquisition and bilingualism can be summarized in terms of inter-individual variability, plasticity and age. The present review looks at variability and plasticity with respect to their underlying sources, and at age as a modulating factor in variability and plasticity. In this context we consider critical period effects vs. bilingualism effects, early and late bilingualism, nativelike and non-nativelike L2 attainment, cognitive aging, individual differences in learning, and linguistic dominance in bilingualism. Non-uniformity is an inherent characteristic of both early and late bilingualism. This review shows how plasticity and age connect with biological and experiential sources of variability, and underscores the value of research that reveals and explains variability. In these ways the review suggests how plasticity, variability and age conspire to frame fundamental research issues in L2 acquisition and bilingualism, and provides points of reference for discussion of the present *Frontiers in Psychology* Research Topic.

Keywords: second language acquisition, bilingualism, plasticity and learning, variability, age factors, individual differences, critical period, dominance

INTRODUCTION

This review article examines a range of features of second-language (L2) acquisition and bilingualism from the intersecting perspectives of plasticity, variability and age. In the simplest terms, for the L2 context plasticity is a property of the neuro-cognitive mechanisms, structures and systems that enable and constrain L2 learning. Variability in L2 attainment at the individual level is conditioned by factors that may be experiential, biological, intellectual, linguistic, conative, educational, and identificational in nature.

Both variability and plasticity are modulated by the age when L2 learning begins [Age of Acquisition (AoA); see below]. Both main and interactive AoA effects on plasticity have been attributed to neurological maturation, to neurochemical and hormonal fluctuations, to decrements of cognitive function over time, to decreases in regional brain volume, to the degree of first-language (L1) entrenchment at the initial state of L2 acquisition, and to the relative use and maintenance of the L1 vs. the L2 (e.g., Birdsong, 2006; Muñoz and Singleton, 2011). AoA may also indirectly condition learner variables such as the extent to which an individual is motivated to acquire an L2 to high levels of proficiency, to engage in the L2 culture, and to identify with L2 speakers (e.g., Dörnyei and Skehan, 2003; Moyer, 2014).

A comprehensive synthesis of relevant research is beyond the scope of this article. Rather, by use of selected examples, the goal is to expose the essential nature of L2 acquisition and bilingualism from the perspectives of age, plasticity and variability. From these perspectives, we can conceive

of linguistic attainment in terms of factors that make L2 learners and bilinguals different from monolinguals, and perhaps get a sense of why these differences are not necessarily deficiencies.

After a brief orientation to the developmental neurobiology of age and plasticity in language learning, I consider the evidence for critical periods in L2 acquisition, taking into account the shape of the function that relates AoA to attainment and the (im)probability of nativelike attainment. The next section offers two illustrations of sources of variation in learning outcomes, which include not only AoA, but also the particular linguistic features under investigation and individuals' cognitive styles and capacities. In the next section I examine possible sources of greater heterogeneity of attainment of L2 morphosyntax with increasing AoA. This is followed by consideration of inter-individual differences, first with respect to exceptional L2 learners and polyglots, then in terms of neurogenetically based talent and trainability, then as a function of idiosyncratic construction of categories for representing linguistic form. In the final section I look at several ways in which linguistic dominance instantiates concerns about plasticity and age, and at how the dominance factor can account for variability in pronunciation and language learning among individual bilinguals.

The works reviewed here converge on the conclusion that studying non-uniformity in language learning outcomes is not so much about sifting through noise and scatter in the data, as it is about illuminating an inherent characteristic of both early and late language acquisition. To this end, it is important to show how plasticity and age connect with biological and experiential sources of variability, and to orient research questions in ways that expose and exploit variability.

A basic motivation of this review is to provide points of reference and theoretical and empirical foundations for readers of the other contributions to the present *Frontiers in Psychology* Research Topic. In so doing I hope to give a sense of how plasticity, variability and age conspire to frame fundamental research issues in L2 acquisition and bilingualism.

Notes on Terminology and Concepts

In this review, the relationship between age and L2 attainment will be considered with respect to the time at which learning of the L2 begins, be it from birth or at any time thereafter. The term AoA refers to the age at which L2 learning begins in earnest and continues with little or no interruption, most often in immersion contexts such as immigration, but not to limited acquaintance with the L2 that takes place in on trips or in the foreign-language classroom. Note that some studies use the terms Age of Exposure, Age of Immersion or Age of Arrival.

The point at which L2 learning begins is conceptualized as the *initial state* of L2 acquisition: the sum of an individual's cognitive, neurological, and linguistic development, along with motivational, identificational, attitudinal and experiential characteristics. Since this cluster of features is difficult to quantify, AoA is taken to be a proxy for the L2 acquisition initial state. In this sense, L2 AoA is understood not as the "age factor" but rather as a "meta-variable" (Flege, 1999). As a predictor variable

in statistical analyses, AoA can be applied to both bilingual (simultaneous or sequential) development in childhood, and to immersion and immigration contexts later in life.

In this review, *bilingualism* is understood to mean routine use of two languages, at whatever level of proficiency in either language. Bilinguals who are immigrants or migrant workers may have acquired their two languages naturalistically only, or they may have had some classroom experience followed by immersion and frequent use. Over the past decade, a disciplinary "bilingual turn" (Ortega, 2009; May, 2014) in language studies recognizes that "L2 learners" and "bilinguals" are not always distinct populations. Obviously, this conflation does not apply to training studies where, for example, participants are taught an artificial language, Mandarin tone contrasts, or the /r/-/l/ distinction in English. Nevertheless, AoA is commonly employed as a predictive factor for learning outcomes in training studies.

In this contribution *critical period* is intended as a generic term that subsumes *sensitive period*. The latter term is sometimes used in contexts of relatively mild maturational effects; at other times it is only meant to suggest heightened receptivity (sensitivity) to relevant environmental stimuli. Both terms refer to finite developmental spans, which may range from birth up to adulthood. In some studies, critical period is taken to mean just the peak period of plasticity or receptiveness of the learning system; in other studies (including the present one), the critical period begins when plasticity starts to increase above baseline and continues until plasticity has leveled out. Maturational effects are thought to take place within, but not beyond, the critical period. For this reason one distinguishes maturational effects from general age effects over the lifespan and, similarly, from AoA effects. For a synopsis of the literature on critical periods for language and other domains, see Birdsong (2017).

Finally, for the purposes of this paper, *learning* and *acquisition* will be used interchangeably. (In some studies, the former term is reserved for formal instructional contexts.)

PLASTICITY, VARIABILITY AND AGE: DEVELOPMENTAL NEUROBIOLOGY AND BEHAVIORAL OUTCOMES

The notion of plasticity with respect to adult language acquisition is often traced back to Penfield and Roberts (1959, p. 240), who argue that for recovery from aphasia the adult brain is "inferior" while the child brain is "plastic," that is, more likely to regain language function. Also seminal in this regard are passing remarks by Lenneberg (1967, p. 176), who links L2 learning difficulties in adulthood with hemispheric functional specialization and declines in plasticity that constrain primary language acquisition.

More recent researchers have put forth other neurobiological explanations for plasticity deficits over age. For example, on a "use it then lose" it model, after adolescence the circuitry that is required for language learning is dismantled because in adulthood there remains no selection pressure on humans to keep learning languages and the metabolically greedy neural systems

that subserve language learning (Hurford, 1991; Pinker, 1994). Another proposed neurobiological culprit is maturationally regulated myelination in the circuitry that underlies language learning. On this view, myelination insulates axons for efficient transmission of electrical impulses, but does so at the cost of reducing the synaptic plasticity required for new learning (Long, 1990; Pulvermüller and Schumann, 1994). Declines in nigrostriatal dopamine over age are implicated in decrements of cognitive abilities such as attention, sequencing, and suppression of competing information; these domain-general capacities are put to use in online L2 processing (Lee, 2004; Wong et al., 2012; see below). The regulation of plasticity takes place within a critical period, "a bounded maturational span during which experiential factors interact with biological mechanisms to determine neurocognitive and behavioral outcomes" (Birdsong, 2017).

To get a fuller sense of the neurobiology of plasticity, and how it might relate to variability in language learning, it is instructive to connect critical-period research in the L1 context with studies in L2 acquisition and bilingualism. The essential neurobiological and experiential characteristics of early language learning are authoritatively laid out by Werker and Hensch (2015), who describe the cascading sequence of multiple, overlapping periods of plasticity that enable the development of phonetic perception in the native language, starting with discrimination of linguistic sounds in the first few months of infancy through the structuring of word forms and phonological categories as children approach 20 months of age; see **Figure 1**.

The chronologies of the onset, the duration, and the closure of each of the critical periods are not fixed, but are manipulated by biological and experiential factors. For example, the timing of the closure of critical periods depends on molecular brakes such as myelin and histone deacetylases, and onset timing can be delayed by sensory deprivation and maternal depression. Thus it is understood that variability and plasticity go hand in hand, as variability within and across overlapping periods of plasticity is a basic feature of the model.

Notably, at the level of the individual child the duration of critical periods in speech perception development can be varied through bilingual experience. As examples, Werker and Hensch cite studies showing that the duration of the critical period for perceptual narrowing – the process by which infants orient their emergent speech perception abilities around just those sounds that occur in their linguistic environment – is longer among simultaneous bilingual children than among monolingual children. The researchers point to several possibilities for this extension.

Relative to monolingual infants, among bilingual infants: native speech categories take longer to establish (Bosch and Sebastián-Gallés, 2003); sensitivities to speech sounds are maintained until an older age (Petitto et al., 2012); there is less input per language, with an asymmetric relative frequency of phones within and across the dual-language input (Bosch and Sebastián-Gallés, 2003); there is enhanced executive control and attentional function afforded by bilingualism (Kovács and Mehler, 2009); the neural circuitry supporting phonetic discrimination is less mature (Garcia-Sierra et al., 2011); the circuitry is equally mature but involves a different distribution of neural connections, with greater connectivity in prefrontal areas (Petitto et al., 2012).

At early developmental stages, the two languages of bilingual infants may resemble those of monolingual children. For example, Burns et al. (2007) found that, at 10–12 months, phonetic discrimination in both languages of English–French bilingual infants of resembled that of monolingual infants and lasted for several months thereafter. Once simultaneous and early bilinguals reach adulthood, however, their processing

FIGURE 1 | Sequential, overlapping critical periods in infant speech perception development. Solid lines represent typical onsets and offsets; broken lines indicate extensions of periods. Adapted from Werker and Hensch (2015). Republished with permission from Annual Reviews.
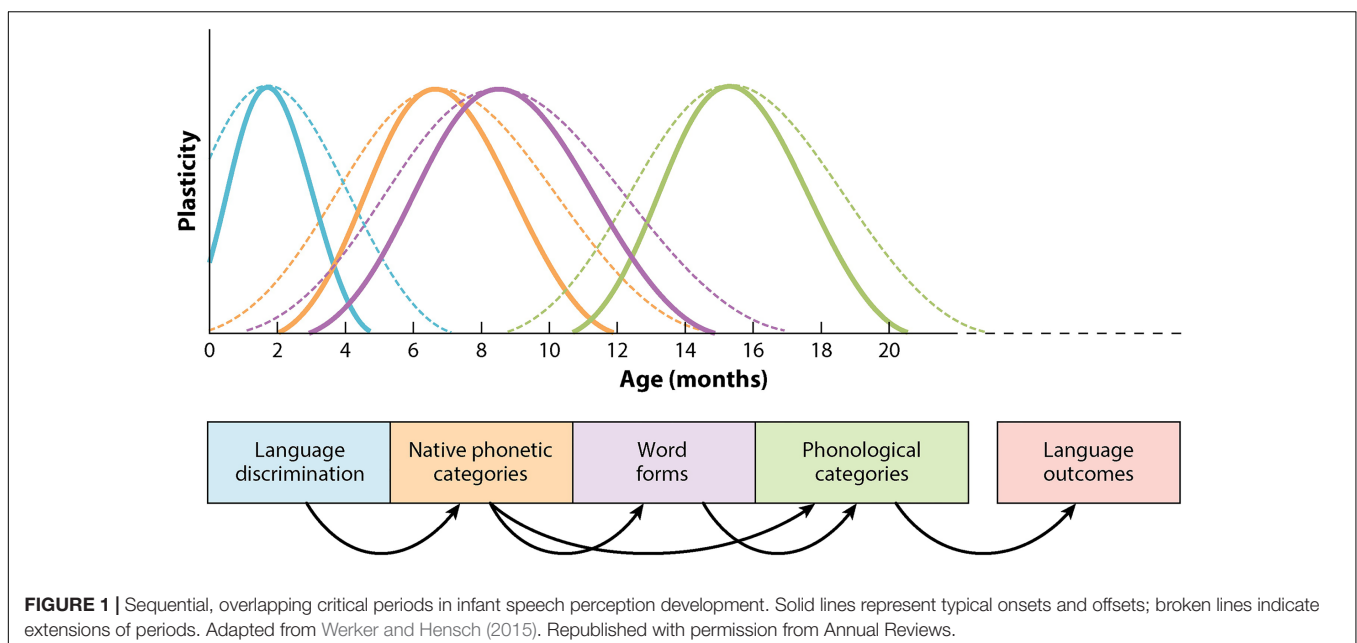
and production of speech differs from that of monolinguals in each language (see below). More to the point, among adult simultaneous and early bilinguals, variability in speech perception and production is widely attested, and the extent of differences among individuals is in general greater than that observed among native (monolingual) speakers; see Sebastián-Galles and Díaz (2012) for a review. This variability (which may reflect asymmetric exposure to or use of the two languages, or exposure to accented speech in one or both languages, along with motivation, context of learning, inter-individual neurobiological and neurocognitive differences over development, etc.; see further discussion below) is often demonstrated in behavioral studies through comparisons of early or simultaneous bilinguals with monolingual controls at local levels of analysis. For example, Mack (1989) looks at early English–French and French–English bilingual adults, all of whom were English dominant. For /ta/-/da/ discrimination and /i/-/I/ production, the bilingual group resembles English monolingual adult controls. In a separate analysis, however, for the percentage of /i/ vowels whose F2 fell at least 50 Hz between the vowel midpoint and offset, bilinguals differ significantly from monolinguals. Similarly, in Sundara et al.'s (2006) study of /d/-/t/ production, English–French simultaneous bilinguals resemble French monolinguals and English monolinguals for /d/ and /t/ in French and for /t/ in English, but diverge for English /d/.

PLASTICITY, VARIABILITY AND CRITICAL PERIODS IN L2 ACQUISITION

It is commonly believed that L2 attainment to nativelike levels among adults is impossible because they have passed a critical period for successful learning. Two general types of evidence are summoned to support this view. The first is the nature of the function that relates AoA to ultimate attainment. The second is evidence for comprehensive nativelike attainment across all aspects of knowledge, production, and processing of the L2.

The AoA-L2 Attainment Function

Theories of the geometry of the function that relates AoA to ultimate (asymptotic) L2 attainment are reviewed in Birdsong (2005) and Birdsong and Vanhove (2016). In brief, it is thought that departures from linearity in the function would suggest effects of developmental events leading to qualitative changes in the neurocognitive mechanisms believed responsible for language learning (see Hakuta et al., 2003, for an overview). If instead the function is linear (**Figure 2A**), this would suggest other types of age-related effects. Some researchers have argued that declines in ultimate L2 attainment should level off after the end of maturation. That is, AoA effects on L2 attainment should be observed among early L2 learners, but AoA should no longer be predictive of L2 asymptote among post-adolescent learners, since maturation would presumably have ceased by this time. On this notion, the geometry of the function should resemble a "stretched L," as seen in **Figure 2B**. On another view, L2 learning is successful up to a certain age (which may vary depending on what language features are being investigated), and learning ability (and, consequently, ultimate attainment) should decline thereafter. The corresponding shape of the function resembles a "stretched 7," as shown in **Figure 2C**. A third geometry is that of a "stretched Z," shown in **Figure 2D**, which combines the "L" and "7" features to include an early plateau, followed by a decline and floor.

To clarify, note that these are schematic representations only. Depending on methodological considerations (e.g., analysis over the AoA span vs. disaggregation by early and late AoA; choice of regression model, line fitting and smoothing methods, etc.) the observed shapes may have less angular features, and the slopes may be shallower. Also, the timing of the points along the AoA continuum where changes in slope are said to occur varies considerably from study to study. (For further discussion, see Birdsong, 2005; Meulman et al., 2015.)

The geometry and timing of AoA effects are crucial to the question of age-conditioned plasticity in L2 learning since, in order to be consistent with maturational effects, the inflection points on the function would need to match up with known

FIGURE 2 | Schematic representations of age of acquisition (AoA) effects on L2 attainment. **(A)** linear decline of L2 attainment over all AoA; **(B)** initial decline of L2 attainment followed by leveling off over subsequent AoA; **(C)** L2 attainment plateau, followed by decline, followed by leveling off over subsequent AoA; **(D)** L2 attainment plateau followed by decline over subsequent AoA.
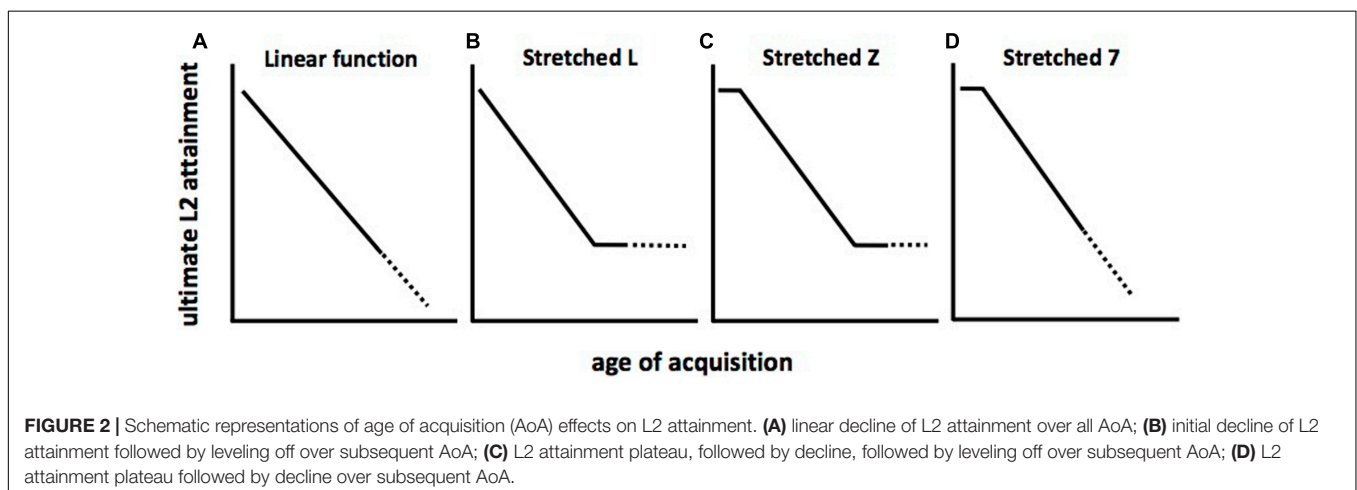
maturational milestones. There are two main obstacles to establishing this isomorphism. One is that attained values on accent ratings and knowledge of morphosyntax map onto different functions. Divergences of L2 learners' pronunciation from that of monolingual controls begin much earlier (even in the first year of life), relative to divergences from native controls for morphosyntax, which have been observed to begin anywhere from about AoA = 7 years to AoA = 27 years. While there is evidence for the intuitively appealing notion of "multiple critical periods" (Scovel, 1988; Granena and Long, 2013), it is a challenge to come up with a unified model that comprehensively aligns variable AoA effects with maturational milestones. Such an account would have to reckon with geometries that are known to vary depending on the pairings of the L1 and the L2, the particular linguistic structures being tested, and exposure, identification, and motivational factors (e.g., Birdsong and Molis, 2001; see further discussion below). In these ways, considerations of plasticity and variability intersect.

Another challenge involves the analytical methods that are employed to generate the AoA- attainment function. Different statistical methods applied to the same data may result in different shapes of the function, thus introducing an additional dimension of variability in our conceptualization of plasticity. For example, in Johnson and Newport's (1989) study of Chinese and Korean learners of L2 English, grammaticality judgment scores on a test of English morphosyntax declined linearly over AoA for learners with AoA $\leq$ 15 years ($r = -0.87$, $p < 0.01$). By contrast, for the later arrivals the scores were distributed more or less randomly ($r = -0.16$, ns), and the best-fitting line through the scatterplot of later-arrivals' scores was roughly horizontal. (Note that this unsystematic dispersion was interpreted by Johnson and Newport as a flattening of the AoA-L2 attainment function; see below). In a subsequent reanalysis of the Johnson and Newport (1989) data, Elman et al. (1996) demonstrated that a single non-linear function accounts for about 63% of the variance over all participants' scores, whereas separate linear regressions for younger and older arrivals account for only about 39% of the variance. Importantly, Elman et al. (1996) point out that the overall best-fitting curve produced by the non-linear model is visually a straight line, i.e., one with no apparent inflection or post-maturational leveling-off.

In a re-examination of the Johnson and Newport (1989) L2 grammaticality judgment data, Vanhove (2013) exposes problems with comparing the correlations for early- vs. late-arriving learners in order to infer maturational effects from different correlational slopes. For Johnson and Newport's early arrivals, the slope of the correlation suggested a decline of scores over AoA, whereas for late arrivals the slope leveled off, with no subsequent AoA-related decline in performance. Together the two correlation slopes resembled a stretched "L" corresponding to one proposed version of a critical period for L2 acquisition. However, as noted above, the apparent "flattened" slope (as indicated by a roughly horizontal regression line) is the reflection of the high degree of variability in the performance of the late-arriving learners.

Vanhove (2013) attributes this essentially random dispersion of late learners' scores to factors such as age-conditioned inter-individual differences in literacy, education, opportunities for L2 use, and motivation – that is, to factors unrelated to critical period constraints. Note as well that general performance levels are often predicted by such variables; see e.g., Birdsong (2014b), Hartshorne et al. (in press).

Vanhove (2013) also reanalyzes L2 grammaticality judgment data from DeKeyser et al. (2010), which involved two groups of Russian native speakers, one having emigrated to Israel and the other to the United States or to Canada. For both participant groups, DeKeyser et al. (2010) had found differences in correlation coefficients between AoA subgroups, and had interpreted the corresponding changes in slope as evidence of discontinuity consistent with critical period effects. In Vanhove's reanalyses, linear and piecewise regressions each account for more than 60% of the variance for both the Israel and North American data. With a breakpoint set at AoA = 18 years, piecewise regressions revealed a linear decline for the Israel data, and only a slight departure from linearity in the North American data.

In contrast, using various regression models, some studies find a "stretched-7" geometry for the AoA-L2 attainment function. For example, Hartshorne et al. (in press) elicited grammaticality judgments for English from 669,498 respondents to an online survey, two-thirds of whom were learners of English from different native languages. The results reveal an L2 ultimate attainment plateau that extends from birth to AoA = 10–12 years, followed by an unbounded decline in judgment accuracy over the remaining AoA range. In a masked priming paradigm involving 94 Turkish–English bilinguals who had learned German at various ages, Veríssimo et al. (2017) observe nativelike priming for inflected German participle forms when the participants' learning began before 5 years of age. After this plateau, facilitation declines with increasing AoA, with no leveling off. Another "stretched-7" geometry is noted by Birdsong and Molis (2001) in their replication study of Johnson and Newport (1989). For 61 native Spanish learners of L2 English, an ultimate attainment plateau terminates at a best-fitting inflection point at AoA = 27.5 years, and performance declines thereafter as AoA increases.

Meulman et al. (2015) illustrate the connectedness of analytical choices, the shape of the AoA-attainment function, and variability across structures under investigation. The researchers looked at ERP P600 signatures for the processing of violations of non-finite verbs and grammatical gender agreement in German by Slavic L1 speakers with advanced proficiency in German L2. AoA effects were not found for non-finite verb violations, which are similar in Slavic and German. However, among participants with AoA $\leq$ 20 years, gender violations elicited a P600, while among those with later AoA a posterior negativity was found in the same time window. Under Generalized Additive Modeling (GAM), and using both AoA and ERP time windows as continuous variables, linear AoA effects on EEG signals were observed across the AoA span, with no discontinuity in the function. Contrarily, ANOVA suggested a critical period prior to AoA = 17.

Non-nativelike Attainment

As a second type of support for critical period effects in L2 acquisition, some researchers point to the lack of evidence for across-the-board nativelikeness in late L2 acquisition (e.g., Long, 1990; Hyltenstam and Abrahamsson, 2003; DeKeyser and Larson-Hall, 2005). The underlying logic is that language learning is biologically destined to be successful if begun in during a critical maturational epoch in early childhood, and that the failure of late learning to attain nativelike competence is the inevitable result of having passed a critical period of neural plasticity. Close comparisons of monolinguals and late L2 learners typically reveal differences across many dimensions of observation (e.g., Abrahamsson and Hyltenstam, 2009), and proponents of the Critical Period Hypothesis for L2 acquisition (CPH/L2A) posit that across-the-board monolingual-likeness is impossible. On this account, in order to falsify the CPH/L2A, one would have to identify at least one late L2 learner who is indistinguishable from a monolingual native across every imaginable measure of linguistic processing and knowledge (Long, 1990).

This argument is implausible, however, because the nature of bilingualism is such that the languages of an active bilingual are activated simultaneously (Dijkstra and Van Heuven, 2002; Schwartz and Kroll, 2006) and influence each other reciprocally (Grosjean, 1989; Cook, 1999, 2003; Flege et al., 2003). Given coactivation and bidirectional effects, neither the first nor the second language of bilinguals can be expected to resemble under scrutiny that of monolinguals in either language. Since "two monolinguals in one person" is an impossibility (Grosjean, 1989), it is unreasonable to hold up a standard of "across-the-board monolingual nativelikeness" in the L2 as a criterion for falsifying the CPH/L2A (Birdsong and Gertken, 2013).

Returning to the question of plasticity, it is important to keep in mind that the L1 is permeable in bilingualism; thus, considerations of plasticity apply to the L1 as well as the L2. The fact that the L2 influences the L1, not just the other way around, suggests that alleged adult L2 learning "deficits" (in the form of divergences from monolingual-likeness) should not be ascribed uniquely to a maturationally determined loss of plasticity. Moreover, the fact that it is not only late L2 learners who exhibit such differences, but early bilinguals and bilinguals-from-birth as well, is plausibly explained under a bilingualism effects account (e.g., MacLeod and Stoel-Gammon, 2005; Fowler et al., 2008; Ortega, 2009). (Note in this context that no researchers claim that bilingualism effects alone are responsible for all divergences from monolingual-likeness in bilingualism.)

Attested non-nativelikeness in both languages of an active bilingual has clear implications for theory. To the extent that an account of L2 acquisition predicts that L2 learners should not attain across-the-board nativelikeness if they have passed a biologically regulated critical period, it should also logically predict that the L1 of a bilingual, which is learned within that critical period, *should* exemplify monolingual-likeness across the board. However, this prediction is not borne out in the relevant research. By contrast, the evidence of bilingualism effects supports an account under which neither the L2 (irrespective of AoA) nor the L1 are completely monolingual-like. Note in this regard that the accuracy figures for bilinguals from birth are significantly lower than those of native monolingual controls in Hartshorne et al. (in press).

These observations connect straightforwardly to questions of age, plasticity, and variability. In their meta-analysis, Liu and Cao (2016) cite studies of L1 permeability in bilingualism, which reveal different patterns of neural activation in the L1 after vs. before acquisition of the L2. Introducing the AoA factor, several reviewed studies converge on the finding that early bilinguals, relative to late bilinguals – with both sampled populations having the same L1 – showed greater activation in the left fusiform gyrus than late bilinguals when processing the L1. This result suggests that the effects of the L2 on L1 processing in imaging studies may be more pronounced with earlier AoA of the L2, as the L2 'interferes' more with the L1 to the extent that development of the two languages overlaps temporally. This relationship is attested as well in behavioral studies.

The basic notion that L2 ultimate attainment is conditioned by the age of initial immersion or significant exposure is examined by Qureshi (2016) in a meta-analysis of 26 studies of morphosyntactic knowledge. The materials reviewed largely substantiated the general idea of AoA effects (as opposed to maturational effects, which were not explicitly examined). At the same time, experiential and methodological factors were found to introduce considerable variability in outcomes. For example, in studies of classroom learning of a foreign language, there was no evidence of an "early advantage" (see also Huang, 2016), whereas the "early-is-better" rule of thumb was supported in studies of immersion learners.

Nativelikeness

It is important to emphasize that, despite bilingualism effects, there are late L2 learners who resemble native monolinguals with respect to targeted aspects of the L2 (as opposed to bilinguals being indistinguishable from monolinguals in every measurable respect). Behavioral evidence ranges from acquisition of fine-grained phonetic features such as VOT to global pronunciation (Bongaerts, 1999; Flege et al., 2002; Birdsong, 2007; Moyer, 2014) and from surface morphology to abstract features of syntax (Birdsong, 1992; Birdsong and Molis, 2001; Donaldson, 2011; Destruel and Donaldson, 2017). In online tasks such as self-paced reading, late bilinguals show monolingual-like sensitivity to subtle and unique aspects of the L2 such as order of clitic pronouns (Rossi et al., 2017). In brain-based studies, high-proficient late L2 learners exhibit convergence with native participants (Green, 2003) in the processing of information structure (Reichle and Birdsong, 2014) and across a variety of syntactic and morpho- syntactic features: see Steinhauer (2014) for a review of the electrophysiological literature; see Abutalebi (2008) for a review of the functional neuroimaging literature.

The incidence of nativelikeness among late L2 learners can vary as a function of the particular structural characteristics that are investigated and as a function of the experimental procedures that are employed. For example, in a series of experiments that involved both ERP and eye-tracking methodologies, Foucart

and Frenck-Mestre (2012) find that violations of noun-adjective gender agreement in French trigger nativelike P600 signatures among English-speaking late learners of L2 French when the adjectives follow the nouns, but elicit non-nativelike N400s when the adjectives are preposed. When the stimuli involve agreement violations in predicative structures (where the noun and the adjective are separated by a copula), natives and learners diverge in terms of ERP, but show similar patterns in eye tracking.

Birdsong and Gertken (2013) point out that the incidence of nativelikeness may depend on which native speakers the learners are being compared to. For example, Indefrey (2006), reviewing studies involving the processing of complex syntax, discerns that natives with high memory spans attend to structural features for correct interpretation in online tasks whereas natives with low memory spans rely on lexico-semantic information – as do many L2 speakers. Indefrey (2006, p. 68) argues that "non-structural sentence processing observed in L2 speakers is an option that is also used by native speakers when they have limited processing resources," thus underscoring another type of variability inherent in assessing nativelikeness.

In some of these and related studies, the findings of nativelikeness have been interpreted as counter-evidence to critical-period predictions with respect to the attainment of nativelikeness in late L2 acquisition articulated. Recall, however, that proponents of the critical period hypothesis in the L2 context advance the criterion of across-the-board nativelikeness as necessary evidence for rejection of the hypothesis. From this perspective, among late (or early) bilinguals it is not enough to find "pockets" of nativelikeness with respect to grammatical knowledge and online processing, or brain activation patterns that resemble those of monolinguals, or individuals who diverge from controls only on VOT values for /d/ in word-final position, but in no other respect.

Under comprehensive, microscopic scrutiny, even among the most practiced hyper-polyglots (see section "Individual Differences in L2 Learning," below), some scintilla of non-monolingual-likeness can be found among active bilinguals.

As stated above, however, the position regarding falsification of the hypothesis by impeccable nativelikeness does not take into account the natural effects of bilingualism, which make it impossible for *both early and late bilinguals* to be exactly like monolinguals *in either the L1 or the L2*. It was also noted that, by the logic of this position, for rejection of the nature-of-bilingualism account (and for support of the critical period account) one would need evidence of across-the-board monolingual-likeness in the first-learned language of late bilinguals, or in either language of simultaneous bilinguals (Birdsong and Vanhove, 2016).

SOURCES OF VARIABILITY IN L2 ACQUISITION

Two Illustrations

Flege et al. (1999) provide an instructive illustration of factors that interact with AoA to produce distinct patterns

of inter-subject variability within the function that relates AoA to L2 attainment. The researchers tested 240 Korean adults' knowledge of L2 English morphosyntax with an adapted version of the Johnson and Newport (1989) materials. **Figure 3A** plots the Koreans' overall performance (black circles) and that of native English controls (open circles). As seen in the plateau at ceiling, participants with early AoA (up to about 7 years of age) perform relatively homogeneously and within or close to the range of native controls.

With increasing AoA, the learners' results become more dispersed. **Figure 3B** plots the L2 English performance on the same items, broken out by those that are grammatical (top image) and ungrammatical (bottom image). Both the top and bottom images reveal increased variability over AoA; however, the degree of variability depends on the grammatical status of the items analyzed, with the cone-shaped scatter of results more pronounced for responses to ungrammatical items than to grammatical items.

Another source of variability is the test items themselves, as shown in **Figure 3C**. For ungrammatical "rule-based" items that exemplify regular, generalizable features of English surface morphology (e.g., –ed past inflection on verbs; case marking on personal pronouns), the slope of the decline in performance over AoA is relatively shallow. By contrast, a steep decline over AoA is observed for ungrammatical "lexically-based" items that exemplify idiosyncratic features of English, such as prepositions preceding infinitival complements (*let to watch vs. let watch) and noun complements (e.g., *hoping rain vs. hoping for rain). Note as well that the shape of the function for the ungrammatical lexical items roughly resembles the schematic "stretched-Z" geometry (see **Figure 2C**), while the function for the rule-based items is closer to linearity (see **Figure 2A**).

As a second illustration of sources of variability, Ettlinger et al. (2014) examine the possibility that L2 learner strategies and success vary according to domain-general cognitive skills. In an artificial language based on Shimakonde, a Bantu language of Mozambique, university student participants were trained on noun stems, plurals, diminutives, and diminutive plurals representing animals. For two types of diminutive plurals in the language, the diminutive and the plural morphemes are simply affixed on the singular stem. A third type of diminutive plurals is more complex, as the vowels in the stem and the plural affix require rephonologization. After exposure to word-picture pairs, participants were asked to produce diminutive plurals on novel words à la the *wug* test (Berko, 1958). Some learners (termed Simplifiers) tended to apply the simple pattern in instances of both complex and simple diminutive plurals; others (Learners) successfully learned both the complex and simple diminutive plurals; others (Non-learners) performed poorly overall. On a prior test of working memory, Learners, Simplifiers and Non-learners performed similarly. However, the groups varied on prior tests of procedural memory and declarative memory. Those participants who were Learners generally scored high on both procedural and declarative memory tests. Those with
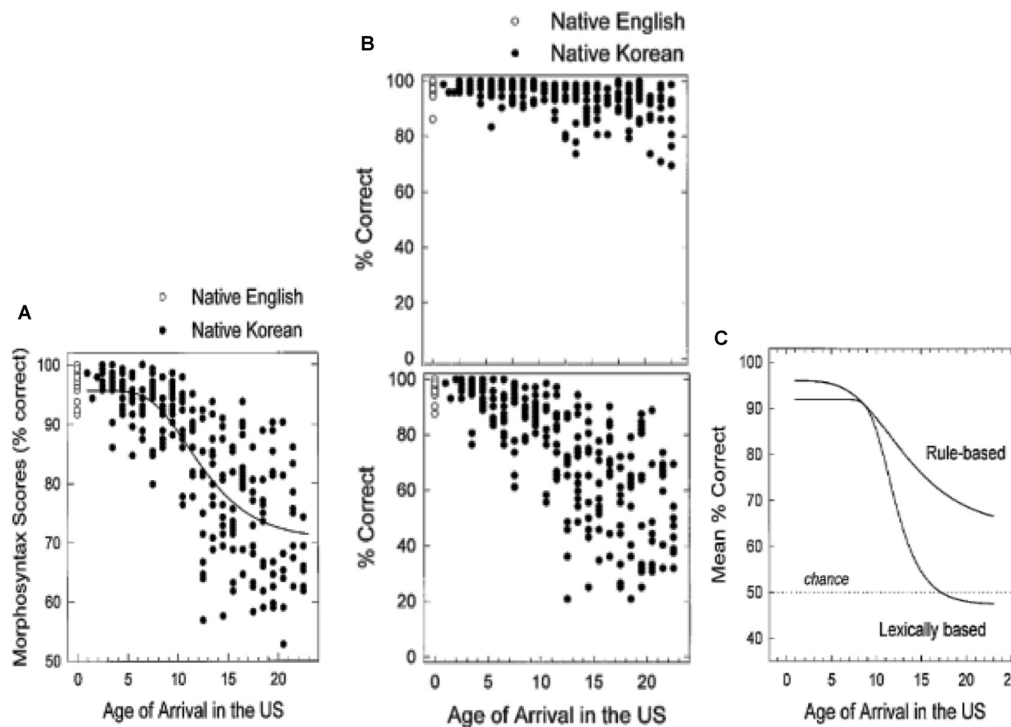
FIGURE 3 | Results of a test of English morphosyntax, as a function of age of arrival in the United States. **(A)** Shows overall percent correct for Korean native speakers (filled circles) and native English controls (open circles); **(B)** breaks out test results by grammatical items (top) and ungrammatical items (bottom); **(C)** depicts different functions for ungrammatical rule-based items vs. ungrammatical lexically-based items. Adapted from Flege et al. (1999). Republished with permission from Elsevier.

high procedural memory scores, but lower declarative memory scores, tended to be Simplifiers. Those with poor procedural memory, irrespective of declarative memory scores, were Non-learners. These results, summarized in **Figure 4**, suggest a connection between learner types and L2 learning performance: differences in domain-general cognitive capacities account for some inter-individual variation in L2 learning.

Variability in L2 Attainment With Increasing AoA: Possible Sources

In some studies, as AoA increases, the outcome of learning of L2 morphosyntax appears to become more variable (see, e.g., Flege et al., 1999; Vanhove, 2013). Candidate sources for such wide dispersions can be inferred from an increase over age of the range of values that are associated with relevant experiential variables. For example, in a random participant sample, the range of lengths of residence in the L2 environment, along with the range of years and types of education will increase correspondingly with AoA. Along with such scaling effects on demographic variables, it is also possible that, with increasing AoA, motivation to attain accuracy in lexico-grammatical knowledge in L2 will become more heterogeneous across participants, particularly so as goals for L2 learning become more diverse.
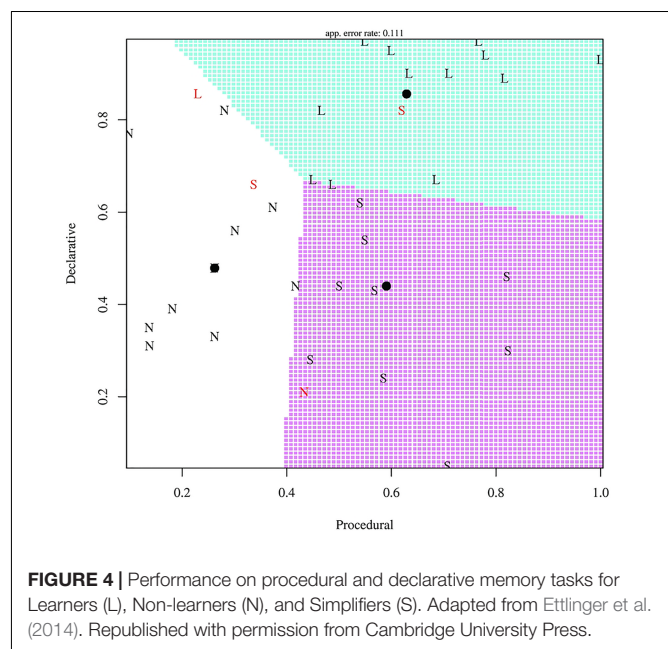
FIGURE 4 | Performance on procedural and declarative memory tasks for Learners (L), Non-learners (N), and Simplifiers (S). Adapted from Ettlinger et al. (2014). Republished with permission from Cambridge University Press.

Cognitive aging may also figure in the mix of candidate reasons for age-related variability in L2 attainment. For example, Buczylowska and Petermann (2016) summarize age-related

differences in six executive function tests administered to 484 participants ranging in age from 18 to 99 years. Declines in mean scores over age were accompanied by increased age-dependent heterogeneity in scores. Connecting this finding to the cone-shaped dispersion of L2 morphosyntax scores over AoA is not a straightforward matter, however, as the heterogeneity observed by Buczylowska and Petermann (2016) is most notable in the later age ranges, whereas most individuals undertaking L2 do not begin so late in life. Further, the degree of dispersion varied greatly by task in this study. Similarly, Mella et al. (2016) show that results on tests of processing speed and working memory do not display the same inter-individual variability with increasing age. Relatedly, Hartshorne and Germine (2015) find that the peaks in cognitive skill are not synchronized over skill types, with some occurring earlier than others. The occurrence of multiple décalages in the timing of peaks (and subsequent declines) suggests that whatever scatter of performance there is over age may not be uniform over intelligence types.

A strong case can be made for both general effects and inter-individual effects of progressive cognitive decline, as well as for effects of dopamine declines (see above), progressive L1 entrenchment (Marchman, 1993; Elman et al., 1996; Flege, 1999; MacWhinney, 2005), and education (Bialystok and Hakuta, 1999; Birdsong, 2014b) on L2 attainment over AoA. At the same time, it is fair to say that further study is needed to establish a direct link between heterogeneity in cognitive function over age and AoA-related patterns of dispersion of results on tests of L2 attainment.

INDIVIDUAL DIFFERENCES IN L2 LEARNING

It is axiomatic that people vary widely in the effectiveness and efficiency with which they learn an L2. Often the study of individual differences in L2 learning focuses on exceptionally successful learners. Although researchers do not all agree on terminological distinctions between the notions of ability, aptitude, talent, and giftedness in the context of L2 learning, the cognitive and conative attributes of high achievers in this domain are well understood; for a recent review, including the question of the mutability of aptitude with experience, see Singleton (2017).

Individuals who attain near-nativelikeness in multiple languages tend to be endowed with high working memory capacity, are highly motivated to learn, and strategically apply metalinguistic knowledge and analysis across their learned languages. In addition to these traits, "gifted multilinguals" score high on tests of intelligence and foreign-language learning aptitude, and are creative, persistent and self-aware (Biedroń and Pawlak, 2016). Polyglots – defined by Hyltenstam (2016) as those who reach high proficiency in six or more languages after puberty – and hyper- polyglots – for Erard (2012) those who proficiently speak, read, or write in at least 11 languages – share the same traits as gifted multilinguals, while also possessing extraordinary verbal memory. They apply their superior analytic skills to recognize patterns in phonology and morphosyntax, and with remarkable executive control are able to switch

between languages with little interference. The linguistic savant Christopher (Smith et al., 2011), who has learned more than 20 languages, exhibits autistic traits and accordingly differs from polyglots and hyperpolyglots in terms of cognitive neurostructure. Pring (2007) notes that autistic savants also differ behaviorally from non-autistic experts by their obsession with memorization and practice, which appears to be more about the pleasure of obsessiveness than about achievement. According to Pring, it is typical of high achievers, but not of savants, to strategically set goals and to use feedback when learning.

Biedroń and Birdsong (in press) point out that, to the extent that complete monolingual nativelikeness is taken to be criterial, extraordinary polyglots do not constitute so-called "exceptions to the critical period hypothesis" for L2 acquisition. As suggested above, it is more apposite to point out that there are no exceptions to the effects of bilingualism, even among the most talented learners of languages. As Biedroń and Birdsong observe, "the special significance of the impossibility of multiple monolingual-likenesses resides in the fact that, no matter how gifted a multilingual is, s/he can't suppress in an absolute sense the other language(s)."

Turning to less exceptional cases, Della Rosa et al. (2013) proffer a view of individual talent in multilingualism that relates language-learning-induced plasticity in the left inferior parietal (LIPL) region of the brain to enhancement of domain-general attentional processes. Their longitudinal study of children living in the South Tyrol region of Italy, where German, Italian, Ladin and English are routinely used, showed specific multilingualism-induced gray matter volume increases in the LIPL. The researchers suggest that such structural adaptations result from the necessity to apply general memory and attentional functions to the processing of more than one language.

A neurogenetic approach to individual differences in L2 learning is advanced by Wong et al. (2012), who specify the mediating roles of genetically encoded dopaminergic (DA) reception and transmission that underlie the acquisition of procedural aspects of grammar. Procedural learning is associated with concatenation of constituents in syntax and with abstract relations between phonology and morphology, and is localized in the prefrontal cortex and basal ganglia. Given what is known about idiosyncratic variability in DA-related gene function, expression and biochemistry, "it is not surprising that individuals with different genetic profiles may have different learning capabilities" (Wong et al., 2012, p. 1093), with more variation expected in adult L2 acquisition than in L1 acquisition. These differences extend to inhibitory function and executive control, which in L2 processing enable suppression of competing information such as knowledge and intrusion of the L1 (Lee, 2004). Under the DA account, a mediating role of AoA can be postulated as well, as dopamine receptor and binding declines over age are well documented (e.g., Volkow et al., 1998; Prull et al., 1999; Bäckman and Farde, 2005).

Taking this approach to variation a step farther, Wong et al. (2017) examine behavioral, neural, and genetic predictors of learning at the level of the individual, and discuss the applications of personalized learning in the L2 context. Drawing parallels with personalized medicine in the pharmacological field, the

authors suggest that understanding individual differences will lead to customization and optimization of language instruction. For other studies of individual differences in cognitive abilities (in particular, differences in procedural, declarative and working memory), and how these play out in second language acquisition, see Morgan-Short et al. (2014) and Faretta-Stutenberg and Morgan-Short (2018).

A dual-systems learning model developed by Chandrasekaran et al. (2014) looks at the use of reflexive vs. reflective learning systems in speech category learning in a training paradigm. The reflective system explicitly develops and tests categorization rules; in contrast, the nature of the reflexive system is procedural and implicit. In experiments involving novel linguistic tone category learning, adult participants initially display a bias toward using the reflective system, which turns out to be ill-adapted to the task. Those individuals who succeed in tone learning are able to shift to the reflexive system, using cortico-striatal connections whose plasticity is regulated by DA reinforcement signals. Relative to younger participants, older adults appear to be less likely to be able to shift from reflective learning to reflexive learning.

Birdsong (2012) examines native-language literacy and education as sources of variability across participants in L2 attainment studies. These factors may interact with task type (e.g., grammaticality judgments vs. truth-value tasks; elicited speech vs. read-alouds), measure (e.g., behavioral vs. brain-based measures; speed vs. accuracy) and linguistic domain (e.g., quantifier scope, garden-path structures). Birdsong (2012) also notes that both native speakers and L2 learners exhibit grammatical idiosyncrasies and other types of variability in representations of linguistic structure (Dabrowska, 2012); therefore variability *per se* (whatever the type or source) is not necessarily evidence of learning deficiencies.

According to Birdsong (1994), the ability to make judgments about linguistic form differs across individuals, who vary in the way they construct language-relevant categories such as "well-formed sentence" and "plausible interpretation." Individual learners may also differ in assessments of the typological relatedness of their L1 to their L2, which modulates their decisions about the likelihood that features of their L1 will resemble those of their L2. Birdsong (2009) characterizes individual differences in learners' ability to notice subtle linguistic features of the L2 within the general framework of signal-detection theory.

For an overview of individual variation in L2 processing (as opposed to attainment), see Van Hell and Abdollahi (2017).

DOMINANCE, PLASTICITY, VARIABILITY AND AGE

A feature of bilingualism that conspicuously connects age, plasticity and variability is linguistic dominance. Regarding plasticity and age, it is not always the case that language learned in infancy is the dominant language of a bilingual: the neural mechanisms involved are sufficiently plastic that the L2 can "leapfrog" the L1 in terms of proficiency and

processing ease. Among international adoptees and heritage speakers, dominance shifts involve attrition of the L1, a representational and functional *loss* which likewise reflects neural plasticity (see below). As concerns variability, inter-individual differences in dominance relationships are natural consequences of idiosyncratic experiences with, skills in, and use of the two languages. No two bilinguals are identical in terms of dominance.

Linguistic dominance in bilingualism is understood in terms of dimensions – relative performance in a language skill such as speech rate, picture naming or grammatical accuracy – and in terms of domains – typically, the comparative frequency of use of each language at work, with family members, or at school. Dominance is not uniquely equatable with relative proficiency (as defined in terms of grammatical and lexical accuracy, speech fluency, etc.), since there are other dimension-based measures of dominance besides proficiency (e.g., object naming speed, lexical diversity, reading speed). Relatedly (and to underscore the dimension/domain distinction in dominance measures), a bilingual parent who is L1-dominant in terms of lexical knowledge and fluency of speech may by choice use the L2 in all interactions with offspring who are being raised in that language, thus demonstrating domain-based L2 dominance in this particular context of use. For further discussion and evidence relating to the independence of dominance and proficiency, see Luk and Bialystok (2013), Montrul (2016a), and Schmeißer et al. (2016); also discussion of balanced bilinguals below.

As with many other features of bilingualism, linguistic dominance is not inherently categorical. That is, individual bilinguals are not simply "L1-dominant" or "L2-dominant," they are dominant in one or the other language to varying degrees. Accordingly, in order to faithfully capture the construct, dominance, like AoA, is properly operationalized and analyzed as a continuous subject factor. As with any other continuous variable, participant assignment to dominance categories may mask intra-group variability and result in loss of statistical power (e.g., Altman, 1998). Some instruments for assessing dominance take into account both domains and dimensions of dominance. Birdsong (2016) reviews methods of calculating dominance indices, along with problems of incommensurability in comparing individual bilinguals who may have the same composite dominance indices, but who vary with respect to the underlying dimensions and domains measured by the instrument.

Balanced Bilingualism

So-called "balanced bilinguals" are dominant in neither language. The term is sometimes used or assumed to denote very high or (near-)nativelike proficiency in both languages. However, degree of proficiency is independent from degree of dominance. An individual who is at an equally low proficiency level in two languages, and an individual who is highly and equally proficient in two languages, are both by definition balanced bilinguals. As depicted by Goto Butler and Hakuta (2004), **Figure 5** shows that balanced bilinguals fall at any point along the diagonal line of increasing proficiency. Bilinguals who are not balanced (that is,

FIGURE 5 | Representation of balanced bilingualism, showing that "balanced" only implies dominance in neither language, not high proficiency in both languages. Adapted from Goto Butler and Hakuta (2004). Republished with permission from John Wiley and Sons.
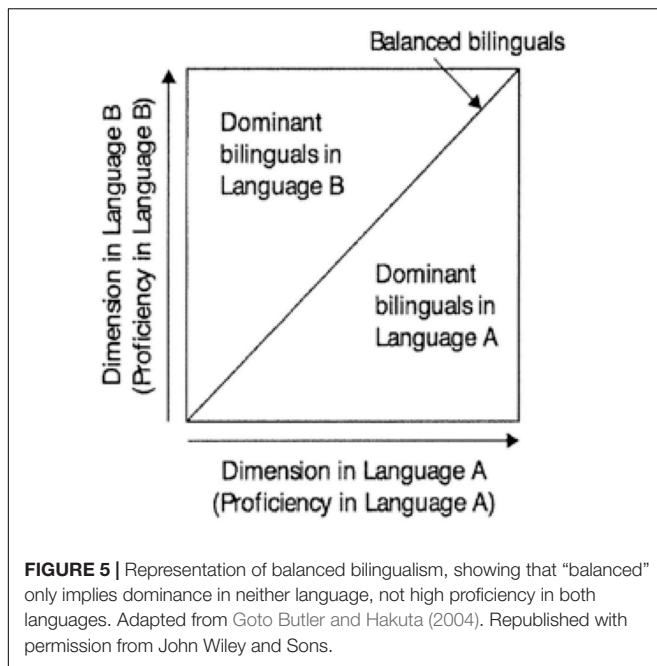
who are dominant in either Language A or Language B) are situated to one side or the other of the diagonal.

The related idea of "perfect bilingualism," if understood as monolingual-likeness in two languages, is misguided since, as noted earlier, neither the L1 nor the L2 of bilinguals is identical to the corresponding languages of monolinguals in all measurable respects. Becoming "more bilingual" is, however, sometimes thought to suggest getting closer to "perfect bilingualism." At the same time "more bilingual" has been taken to mean that a given bilingual is highly proficient in the two languages, or to mean approaching balanced proficiency.

Dominance Shifts and Age

The direction and degree of dominance in the two languages are dynamic over the lifetime of an individual bilingual. Depending on changing circumstances such as immigration, educational and occupational opportunities, and psycho-social identification, the L2 may "replace" the L1 as the dominant language. In some cases, and for similar reasons, the L1 may return to dominance, and still further shifts are possible. Grosjean (2010) details multiple dominance shifts over 60 years of his life. For a review of research and theory on the relationship between dominance and age, see Birdsong (2014a).

Conceptually as well as in practice, the developmental dynamics of dominance relationships may reflect both L1 loss and L2 gains. For example, among some immigrants and adoptees, there may be little or no ongoing use of the L1; as the L1 withers (in terms of domains or dimensions), the L2 perforce becomes the dominant language. On a developmental scenario, a sequential bilingual whose L1 is not fully developed may use and maintain the L1, but as a matter of relative gains in linguistic knowledge and proficiency over time, the L2 eventually outstrips the L1.

Losses in the L1 and gains in the L2, with consequent reflexes in the dominance relationship between the two languages, have been theorized together in terms of maturational constraints on plasticity. Bylund et al. (2012) propose that the same maturational mechanisms synchronously constrain both the ability to lose a language and the ability to gain a new language. Bylund et al. (2012) state that the potential for L1 attrition and the potential for L2 attainment are highest during the first 10 years of life. After this period the potential for both L1 attrition and L2 attainment declines, with the relevant geometry of both resembling a stretched-7. Pallier (2007) advances a different view, whereby the AoA-ultimate L2 attainment function exhibits a linear decline, starting essentially at birth; by contrast, the likelihood and degree of L1 attrition start to drop off only after age 10. For Pallier (2007) and Bylund et al. (2012) alike, plasticity for both L1 attrition and L2 attainment are age conditioned; however, for Pallier the age effects for L2 attainment do not correspond to maturational effects in the AoA-attainment function, as there is no departure from linearity along the function that would suggest a qualitative change in learning ability.

Thus, with respect to plasticity in dominance relationships in the first decade of life, there are two distinct possibilities. One possibility is that L1 loss, the likelihood of which is highest for several early years, is a greater contributor to dominance shifts than L2 gains, which start to become less likely very early in life, with progressively less influence on shifts from L1 to L2 dominance. Another possibility is that L2 gains and L1 losses conspire simultaneously to enable L1-to-L2 shifts of dominance. The latter possibility relates L1 loss and L2 gain under a unified view of plasticity in early childhood development: "the ease with which an L2 is acquired and the L1 undergoes attrition can be said to be manifestations of a generally heightened responsiveness to language exposure, which works both in acquisitional and attritional directions" (Bylund et al., 2012, p. 237). For a recent empirical study and review of age effects on L1 attrition, see Ahn et al. (2017).

Note that age conditions not only the probability of L1 loss, but also the speed at which attrition occurs (Köpke and Schmid, 2004). As L1 loss slows, the point at which a complete shift to L2 dominance can be expected is delayed. Similarly, depth of attrition (the degree to which a domain or dimension is diminished) and breadth of attrition (the number of dimensions and domains diminished) should decrease with the age at which the loss begins. Thus, indirectly through L1 loss, age contributes to variability in L1–L2 dominance relationships (see also Montrul, 2016b).

Examples of Prediction and Variation in Dominance

Dominance has been shown to be a predictive factor in studies of bilingualism. As an example, Amengual (2014) looks at the elicited production of mid vowels among Spanish–Catalan bilinguals in Majorca. Catalan, but not Spanish, makes a phonemic distinction between the tense-lax mid /e/ - /ɛ/ and /o/ - /ɔ/. Relative Catalan vs. Spanish dominance was assessed with the Bilingual Language Profile (BLP; Birdsong et al., 2012). For

the 30 Catalan-dominant bilinguals, degree of dominance was not predictive of the Euclidian distance between /e/ and /ɛ/ nor between /o/ and /ɔ/. However, among the 30 Spanish dominants, those whose BLP scores approached balanced bilingualism (i.e., whose scores were least Spanish dominant) produced the vowels with Euclidean distances resembling those produced by the Catalan dominants. Specifically, the BLP scores of Spanish dominants were predictive of more Catalan-like Euclidean distance between /e/ and /ɛ/ and between /o/ and /ɔ/ (see **Figure 6**).

A study of bilingual speakers in Guatemala by Baird (2015) illustrates how the dominance factor accounts for inter-individual variation in bilingualism. Baird examines the pronunciation of Spanish tonic syllables by Spanish–K'ichee' bilinguals from two Guatemalan communities, Cantel and Nahualá. In most varieties of Spanish, the peak of F0 rise occurs after the tonic syllable. In contact and bilingualism contexts, Spanish varieties display an F0 that is closer to (sometimes before) the tonic syllable. In a task involving reading Spanish phrases, 10 Spanish monolinguals, 10 bilinguals from Cantel, and 7 of the 10

bilinguals from Nahualá produced late (post-stress) F0 peaks. At the same time, for speakers from both communities, the degree of Spanish vs. K'ichee' dominance, as assessed by the BLP, was predictive of the direction and distance of F0 peak placement; see **Figure 7** (Baird, 2015).

A critical take-away from Baird (2015) is that the nature of inter-individual variation is obscured in a simple analysis by binary factors, in this instance place of residence and pre- vs. post-stress F0 peaks. More revealing can be examinations of variation along continuous dimensions, in this case distance of peaks from the tonic syllable and degree of Spanish vs. K'ichée dominance. By such an analysis, individual variability along a continuum of peak F0 placement is predicted by degree of dominance, independently of residence.

Researchers have considered the possibility that dominance in the L2 may be associated with monolingual nativelikeness in pronunciation in that language. In a delayed sentence-repetition task for English sentences, Flege et al. (2002) found that Italian–English bilinguals who were L2-English dominant were judged not to have foreign accents, and suggested that
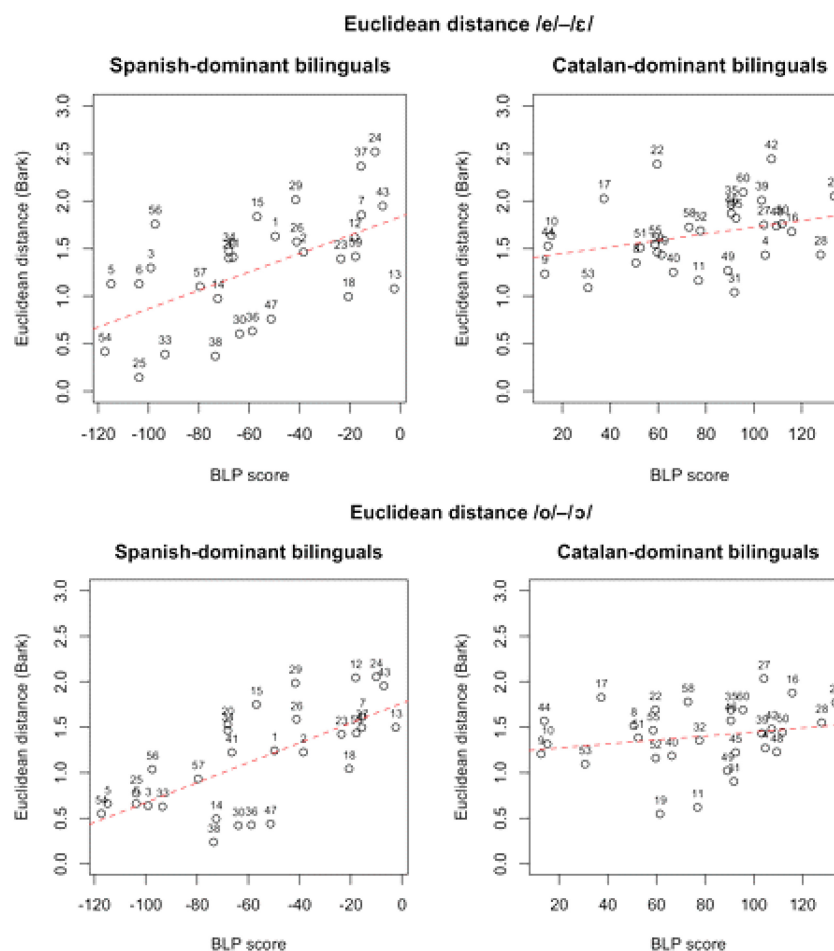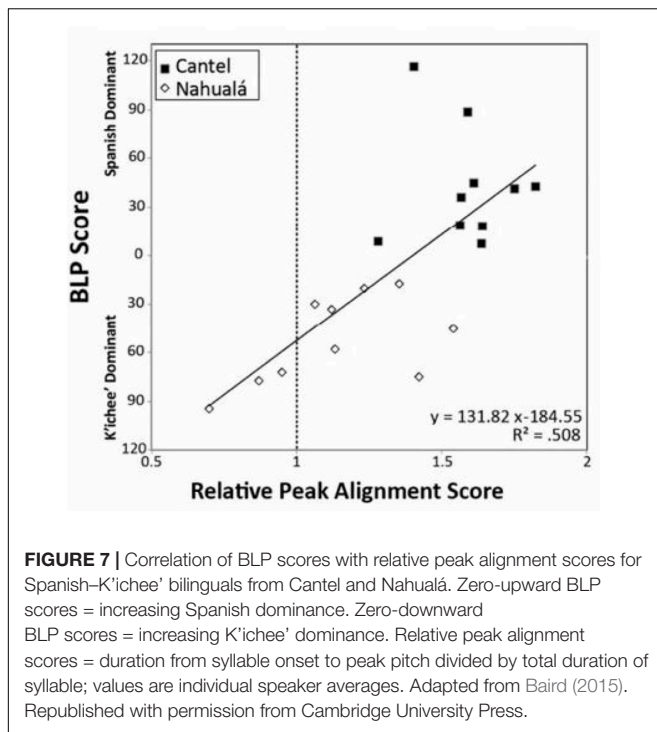

FIGURE 6 | For Spanish-dominant and Catalan-dominant Majorcan bilinguals, Euclidean distances between same-speaker tokens of /e/-/ɛ/ and /o/-/ɔ/, plotted as a function of BLP scores, which range from −120 (strongest Spanish dominance) to +130 (strongest Catalan dominance). Adapted from Amengual (2014). Republished with permission from Sage Publishing.

FIGURE 7 | Correlation of BLP scores with relative peak alignment scores for Spanish–K'ichee' bilinguals from Cantel and Nahualá. Zero-upward BLP scores = increasing Spanish dominance. Zero-downward BLP scores = increasing K'ichee' dominance. Relative peak alignment scores = duration from syllable onset to peak pitch divided by total duration of syllable; values are individual speaker averages. Adapted from Baird (2015). Republished with permission from Cambridge University Press.

L1-interference effects might be absent among L2-dominant bilinguals. A series of follow-up studies by Antoniou and colleagues look more closely at interference effects, in this case with respect to VOT among L2 dominants. In a sample of Greek–English sequential bilinguals who were L2 English-dominant, Antoniou et al. (2010) find that stop voicing among the L2-dominants mostly match that of natives in both languages, with exceptions for some L2 English medial stops reflecting a measure of L1 Greek interference. For the same bilinguals, Antoniou et al. (2011) examine VOT in code-switching between English and Greek. In contrast to the unilingual mode (one language activated) results of Antoniou et al. (2010), English stops in bilingual mode (both languages activated) are produced with more Greek-like values, whereas Greek stops do not display English-like VOT. That is, the L1 appears to influence pronunciation in the dominant L2, but not the other way around. Perception experiments with a larger sample of Greek–English bilinguals (Antoniou et al., 2012, p. 592) reveal a still more complex pattern of dominance relationships, one that depends on whether the task is categorization or discrimination of voicing: "The results suggest that a bilingual is a single (dominant-language) listener with respect to discrimination, but behaves more like a monolingual of the activated language with respect to discrimination judgments." Taken together, the findings of Flege and colleagues and those of Antoniou and colleagues suggest a high degree of variability in terms of monolingual-like performance among L2-dominant bilinguals. Results may vary according to production vs. perception, language mode (unilingual vs. bilingual), task (discrimination vs. categorization), and level of analysis (global pronunciation vs. VOT).

Another illustration of the role of dominance in bilingualism relates to the question of executive control. A considerable body of research (e.g., Bialystok et al., 2012) suggests that enhanced executive control is conferred by bilingualism. At the same time, since bilingualism is not a unitary phenomenon and thus not a categorical variable (Luk and Bialystok, 2013); Yow and Li (2015) examine degree of dominance as a predictor of cognitive control within bilingual populations. Among 72 English–Mandarin young adult bilinguals, the researchers find a positive effect for balanced use and balanced proficiency with respect to interference in Stroop task performance and mixing cost in a number-letter (mental-set shifting) task. In addition, early AoA of the second language is associated with less interference on the Stroop task.

As a related and final example, recent work by Onnis et al. (2018) looks at dominance as a predictor of statistical learning among adult bilinguals in a miniature grammar paradigm. (Statistical language learning involves tracking the frequencies of, or the transitional probabilities between, grammatical elements, which results in implicit knowledge of structural regularities.) In this study, success in statistical learning of artificial grammars is predicted by the degree to which participants approach or depart from balanced bilingualism, as measured by BLP scores: balanced bilinguals perform better than those who are increasingly dominant in their first language. Thus, degree of bilingual dominance in adulthood is associated with differential ability to learn a novel language. Onnis et al. (2018, p. 432) summarize their findings: "By capitalizing on the bilingual variability we found in the [BLP] questionnaire rather than ignoring it, we unearthed important individual differences that point to the first documented modulating role of [degree of dominance in] bilingualism in adult statistical learning."

CONCLUSION

In this review we have seen how variation in L2 acquisition and bilingualism is conditioned by age, which itself conditions plasticity. We also know that age similarly conditions individual factors such as language experience, L1 attrition and linguistic dominance, which are themselves predictive of variation.

Age-related effects (of which neurobiological maturation within a critical period is one possible source) cannot account for all varieties of non-nativelike outcomes in L2 acquisition, since departures from monolingual-likeness are found not just in post-childhood learning but among from-birth simultaneous bilinguals as well. By contrast, bilingualism effects can account for observed non-monolingual-likeness in both the L1 and the L2, whatever the age of learning. At the same time, the degree of L1 activation, L1 entrenchment, L1 attrition and relative L1–L2 dominance – all of which are affected by AoA – modulate attainment levels across L2 learners.

The application of different statistical models and methods can result in different shapes of the function that relates AoA to L2 outcomes; such artifacts add another dimension of variability

to the picture of L2 acquisition. We have also considered possible sources of variability in L2 attainment with increasing AoA. These sources range from experiential (education, length of residence), to representational (L1 entrenchment) and to cognitive decline with underlying neurologic causes such as dopamine levels that mediate domain-general learning and processing. The role of cognitive decline in AoA-related variability in L2 learning outcomes is of particular interest for future investigation.

This review has brought these concerns into focus with illustrations from two areas of active research, individual differences and bilingual dominance. With respect to individual differences in L2 learning, we have highlighted the roles of neurogenetic makeup, higher-order cognitive factors, language experience, age-conditioned learning styles and motivation. We have seen that the gradient phenomenon of dominance in bilingualism is dynamic over the lifespan, is conditioned by experience as well as by neural plasticity, and is predictive of

phonetic variation, cognitive control, and statistical learning in artificial language paradigms.

In his classic position paper Bley-Vroman (1990, p. 13) problematizes adult L2 learning in terms of explaining "the quite high level of competence that is clearly possible in some cases, while also permitting the wide range of variation that is observed." By demonstrating the connectedness of non-uniform outcomes with age and plasticity, the research reviewed here has shown that such variation is neither unexplainable nor unexpected. From this understanding emerges heuristic guidance for further explorations of the richness of L2 acquisition and bilingualism.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and approved it for publication.

REFERENCES

Abrahamsson, N., and Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: listener perception versus linguistic scrutiny. *Lang. Learn.* 59, 249–306. doi: 10.1111/j.1467-9922.2009.00507.x

Abutalebi, J. (2008). Neural aspects of second language representation and language control. *Acta Psychol.* 128, 466–478. doi: 10.1016/j.actpsy.2008.03.014

Ahn, S., Chang, C. B., DeKeyser, R., and Lee-Ellis, S. (2017). Age effects in first language attrition: speech perception by Korean-English bilinguals. *Lang. Learn.* 67, 694–733. doi: 10.1111/lang.12252

Altman, D. G. (1998). "Categorizing continuous variables," in *Encyclopedia of Biostatistics*, eds P. Armitage and T. Colton (New York, NY: Wiley), 563–567.

Amengual, M. (2014). The perception and production of language-specific mid-vowel contrasts: shifting the focus to the bilingual individual in early language input conditions. *Int. J. Biling.* 20, 133–152. doi: 10.1177/1367006914544988

Antoniou, M., Best, C. T., Tyler, M. D., and Kroos, C. (2010). Language context elicits native-like stop voicing in early bilinguals' productions in both L1 and L2. *J. Phon.* 38, 640–653. doi: 10.1016/j.wocn.2010.09.005

Antoniou, M., Best, C. T., Tyler, M. D., and Kroos, C. (2011). Inter-language interference in VOT production by L2-dominant bilinguals: asymmetries in phonetic code-switching. *J. Phon.* 39, 558–570. doi: 10.1016/j.wocn.2011.03.001

Antoniou, M., Tyler, M. D., and Best, C. T. (2012). Two ways to listen: do L2-dominant bilinguals perceive stop voicing according to language mode? *J. Phon.* 40, 582–594. doi: 10.1016/j.wocn.2012.05.005

Bäckman, L., and Farde, L. (2005). "The role of dopamine systems in cognitive aging," in *Cognitive Neuroscience of Aging: Linking Cognitive and Cerebral Aging*, eds R. Cabeza, L. Nyberg, and D. Park (New York, NY: Oxford University Press), 58–84.

Baird, B. O. (2015). "Pre-nuclear peak alignment in the Spanish of Spanish-K'ichee' (Mayan) bilinguals," in *Proceedings of the 6th Conference on Laboratory Approaches to Romance Phonology*, eds E. W. Willis, P. Martín Burtagueño, and E. Herrera Zendejas (Somerville, MA: Cascadilla Proceedings Project), 163–174.

Berko, J. (1958). The child's learning of English morphology. *Word* 14, 150–177. doi: 10.1080/00437956.1958.11659661

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., and Hakuta, K. (1999). "Confounded age: linguistic and cognitive factors in age differences for second language acquisition," in *Second Language Acquisition and the Critical Period Hypothesis*, ed. D. Birdsong (Mahwah, NJ: Lawrence Erlbaum Associates), 161–181.

Biedroń, A., and Birdsong, D. (in press). "Highly proficient and gifted bilinguals," in *Handbook of Bilingualism*, eds L. Ortega and A. De Houwer (Cambridge: Cambridge University Press).

Biedroń, A., and Pawlak, M. (2016). New conceptualizations of linguistic giftedness. *Lang. Teach.* 49, 151–185. doi: 10.1017/S026144481500039

Birdsong, D. (1992). Ultimate attainment in second language acquisition. *Language* 68, 706–755. doi: 10.1353/lan.1992.0035

Birdsong, D. (1994). Decision making in second language acquisition. *Stud. Second Lang. Acquis.* 16, 169–182. doi: 10.1017/S0272263100012869

Birdsong, D. (2005). "Interpreting age effects in second language acquisition," in *Handbook of Bilingualism: Psycholinguistic Approaches*, eds J. F. Kroll and A. M. B. de Groot (New York, NY: Oxford University Press), 109–127.

Birdsong, D. (2006). Age and second language acquisition and processing: a selective overview. *Lang. Learn.* 56, 9–49. doi: 10.1111/j.1467-9933.2006.00353.x

Birdsong, D. (2007). "Nativelike pronunciation among late learners of French as a second language," in *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*, eds O.-S. Bohn and M. J. Munro (Amsterdam: John Benjamins), 99–116.

Birdsong, D. (2009). Uninterpretable features: psychology and plasticity in second language learnability. *Second Lang. Res.* 25, 235–243. doi: 10.1177/0267658308100285

Birdsong, D. (2012). Three perspectives on non-uniform linguistic attainment. *Linguist. Approaches Biling.* 2, 255–259. doi: 10.1075/lab.2.3.02bir

Birdsong, D. (2014a). Dominance and age in bilingualism. *Appl. Linguist.* 35, 374–392. doi: 10.1093/applin/amu031

Birdsong, D. (2014b). "The critical period hypothesis for second language acquisition: tailoring the coat of many colors," in *Essential Topics in Applied Linguistics and Multilingualism: Studies in Honor of David Singleton*, eds M. Pawlak and L. Aronin (New York, NY: Springer), 43–50.

Birdsong, D. (2016). "Dominance in bilingualism: foundations of measurement, with insights from the study of handedness," in *Language Dominance in Bilinguals: Issues of Measurement and Operationalization*, eds C. Silva-Corvalán and J. Treffers-Daller (Cambridge: Cambridge University Press), 85–105. doi: 10.1017/CBO9781107375345.005

Birdsong, D. (2017). "Critical periods," in *Oxford Bibliographies in Linguistics*, ed. M. Aronoff (New York, NY: Oxford University Press).

Birdsong, D., Gertken, L. M., and Amengual, M. (2012). *Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism*. Austin, TX: COERLL, University of Texas at Austin. Available at: https://sites.la.utexas.edu/bilingual/

Birdsong, D., and Gertken, L. M. (2013). In faint praise of folly: a critical review of native/non-native comparisons, with examples from native and bilingual processing of French complex syntax. *Lang. Interact. Acquis.* 4, 107–133. doi: 10.1075/lia.4.2.01bir

Birdsong, D., and Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *J. Mem. Lang.* 44, 235–249. doi: 10.1006/jmla.2000.2750

Birdsong, D., and Vanhove, J. (2016). "Age of second-language acquisition: critical periods and social concerns," in *Bilingualism Across the Lifespan: Factors Moderating Language Proficiency*, eds E. Nicoladis and S. Montanari (Washington, DC: American Psychological Association), 162–181.

Bley-Vroman, R. (1990). The logical problem of foreign language learning. *Linguist. Anal.* 20, 3–50. doi: 10.1017/CBO9781139524544.005

Bongaerts, T. (1999). "Ultimate attainment in L2 pronunciation: the case of very advanced late learners," in *Second Language Acquisition and the Critical Period Hypothesis*, ed. D. Birdsong (Mahwah, NJ: Lawrence Erlbaum Associates), 133–159.

Bosch, L., and Sebastián-Gallés, N. (2003). Simultaneous bilingualism and perception of a language-specific vowel contrast in the first year of life. *Lang. Speech* 46, 217–243. doi: 10.1177/00238309030460020801

Buczylowska, D., and Petermann, F. (2016). Age-related differences and heterogeneity in executive functions: analysis of NAB executive functions module scores. *Arch. Clin. Neuropsychol.* 31, 254–282. doi: 10.1093/arclin/acw005

Burns, T. C., Yoshida, K. A., Hill, K. M., and Werker, J. F. (2007). The development of phonetic representation in bilingual and monolingual infants. *Appl. Psycholinguist.* 28, 455–474. doi: 10.1017/S0142716407070257

Bylund, E., Abrahamsson, N., and Hyltenstam, K. (2012). Does first language maintenance hamper nativelikeness in a second language? *Stud. Second Lang. Acquis.* 34, 215–241. doi: 10.1017/S0272263112000034

Chandrasekaran, B., Koslov, S. R., and Maddox, W. T. (2014). Toward a dual-learning systems model of speech category learning. *Front. Psychol.* 5:825. doi: 10.3389/fpsyg.2014.00825

Cook, V. (1999). Going beyond the native speaker in language teaching. *TESOL Q.* 33, 185–209. doi: 10.2307/3587717

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

Dabrowska, E. (2012). Different speakers, different grammars: individual differences in native language attainment. *Linguist. Approaches Biling.* 2, 219–253. doi: 10.1075/lab.2.301dab

DeKeyser, R., Alfi-Shabtay, I., and Ravid, D. (2010). Cross-linguistic evidence for the nature of age effects in second language acquisition. *Appl. Psycholinguist.* 31, 413–438. doi: 10.1017/S0142716410000056

DeKeyser, R., and Larson-Hall, J. (2005). "What does the critical period really mean?," in *Handbook of Bilingualism: Psycholinguistic Approaches*, eds J. F. Kroll and A. M. B. de Groot (New York, NY: Oxford University Press), 88–108.

Della Rosa, P. A., Videsott, G., Borsa, V. M., Canini, M., Weekes, B. S., Franceschini, R., et al. (2013). A neural interactive location for multilingual talent. *Cortex* 49, 605–608. doi: 10.1016/j.cortex.2012.12.001

Destruel, E., and Donaldson, B. (2017). Second language acquisition of pragmatic inferences: evidence from the French *c'est*-cleft. *Appl. Psycholinguist.* 38, 703–732. doi: 10.1017/S014271614600400

Dijkstra, A., and Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: from identification to decision. *Bilingualism* 5, 175–197. doi: 10.1017/S1366728902003012

Donaldson, B. (2011). Left-dislocation in near-native French. *Stud. Second Lang. Acquis.* 33, 399–432. doi: 10.1017/S0272263111000039

Dörnyei, Z., and Skehan, P. (2003). "Individual differences in second language learning," in *The Handbook of Second Language Acquisition*, eds C. J. Doughty and M. H. Long (Malden, MA: Blackwell), 589–630. doi: 10.1002/9780470756492.ch18

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., and Plunkett, K. (1996). *Rethinking Innateness: A Connectionist Perspective on Development*. Cambridge, MA: MIT Press.

Erard, M. (2012). *Babel No More: The Search for the World's Most Extraordinary Language Learners*. New York, NY: Free Press.

Ettlinger, M., Bradlow, A., and Wong, P. C. M. (2014). Variability in the learning of complex morphophonology. *Appl. Psycholinguist.* 35, 807–831. doi: 10.1017/S0142716412000586

Faretta-Stutenberg, M., and Morgan-Short, K. (2018). The interplay of individual differences and context of learning in behavioral and neurocognitive second language development. *Second Lang. Res.* 34, 67–101. doi: 10.1177/0267658316684903

Flege, J. E. (1999). "Age of learning and second language speech," in *Second Language Acquisition and the Critical Period Hypothesis*, ed. D. Birdsong (Mahwah, NJ: Lawrence Erlbaum Associates), 101–131.

Flege, J. E., MacKay, I. R. A., and Piske, T. (2002). Assessing bilingual dominance. *Appl. Psycholinguist.* 23, 567–598. doi: 10.1017/S0142716402004046

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

Flege, J. E., Yeni-Komshian, G. H., and Liu, S. (1999). Age constraints on second-language acquisition. *J. Mem. Lang.* 41, 78–104. doi: 10.1006/jmla.1999.2638

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

Fowler, C. A., Sramko, V., Ostry, D. J., Rowland, S. A., and Hallé, P. (2008). Cross language phonetic influences on the speech of French-English bilinguals. *J. Phon.* 36, 649–663. doi: 10.1016/j.wocn.2008.04.001

Garcia-Sierra, A., Rivera-Gaxiola, M., Peraccio, C. R., Conboy, B. T., Romo, H., Klarman, S. O., et al. (2011). Bilingual language learning: an ERP study relating early brain responses to speech, language input, and later word production. *J. Phon.* 39, 546–557. doi: 10.1016/j.wocn.2011.07.002

Goto Butler, Y., and Hakuta, K. (2004). "Bilingualism and second language acquisition," in *Handbook of Bilingualism*, eds T. Bhatia and W. Ritchie (New York: Blackwell Publishers), 114–144.

Granena, G., and Long, M. H. (2013). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Lang. Res.* 29, 311–343. doi: 10.1177/0267658312461497

Green, D. W. (2003). "The neural basis of the lexicon and the grammar in L2 acquisition: the convergence hypothesis," in *The Interface Between Syntax and the Lexicon in Second Language Acquisition*, eds R. van Hout, A. Hulk, F. Kuiken, and R. Towell (Amsterdam: John Benjamins), 197–218. doi: 10.1075/lald.30.10gre

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain Lang.* 36, 3–15. doi: 10.1016/0093-934X(89)90048-5

Grosjean, F. (2010). *Bilingual: Life and Reality*. Cambridge, MA: Harvard University Press. doi: 10.4159/9780674056459

Hakuta, K., Bialystok, E., and Wiley, E. W. (2003). Critical evidence: a test of the critical-period hypothesis for second-language acquisition. *Psychol. Sci.* 14, 31–38. doi: 10.1111/1467-9280.01415

Hartshorne, J. K., and Germine, L. T. (2015). When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the life span. *Psychol. Sci.* 26, 433–443. doi: 10.1177/0956797614567339

Hartshorne, J. K., Tenenbaum, J. B., and Pinker, S. (in press). A critical period for second language acquisition: evidence from 2/3 million English speakers. *Cognition*.

Huang, B. H. (2016). A synthesis of empirical research on the linguistic outcomes of early foreign language instruction. *Int. J. Multiling.* 13, 257–273. doi: 10.1080/14790718.2015.1066792

Hurford, J. R. (1991). The evolution of the critical period for language acquisition. *Cognition* 40, 159–201. doi: 10.1016/0010-0277(91)90024-X

Hyltenstam, K. (ed.). (2016). *Advanced Proficiency and Exceptional Ability in Second Languages*. Berlin: Mouton de Gruyter. doi: 10.1515/9781614515173

Hyltenstam, K., and Abrahamsson, N. (2003). "Maturational constraints in SLA," in *The Handbook of Second Language Acquisition*, eds C. J. Doughty and M. H. Long (Malden, MA: Blackwell), 539–588.

Indefrey, P. (2006). It is time to work toward explicit processing models for native and second language speakers. *Appl. Psycholinguist.* 27, 66–69. doi: 10.1017/S0142716406280032

Johnson, J. S., and Newport, E. L. (1989). Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a Second Language. *Cogn. Psychol.* 21, 60–99. doi: 10.1016/0010-0285(89)90003-0

Köpke, B., and Schmid, M. S. (2004). "Language attrition: the next phase," in *First Language Attrition: Interdisciplinary Perspectives on Methodological Issues*, eds M. S. Schmid, B. Köpke, M. Keijzer, and L. Weilemar (Amsterdam: John Benjamins), 1–43.

Kovács, Á. M., 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

Lee, N. (2004). "The neurobiology of procedural memory," in *The Neurobiology of Learning: Perspectives from Second Language Acquisition*, eds J. H. Schumann, S. E. Crowell, N. E. Jones, N. Lee, S. A. Schuchert, and L. A. Wood (Mahwah, NJ: Lawrence Erlbaum Associates), 43–73.

Lenneberg, E. H. (1967). *Biological Foundations of Language*. New York, NY: Wiley.

Liu, H., and Cao, F. (2016). L1 and L2 processing in the bilingual brain: a meta-analysis of neuroimaging studies. *Brain Lang.* 159, 60–73. doi: 10.1016/j.bandl. 2016.05.013

Long, M. H. (1990). Maturational constraints on language development. *Stud. Second Lang. Acquis.* 12, 251–285. doi: 10.1017/S0272263100009165

Luk, G., and Bialystok, E. (2013). Bilingualism is not a categorical variable: interaction between language proficiency and usage. *J. Cogn. Psychol.* 25, 605–621. doi: 10.1080/20445911.2013.759574

Mack, M. (1989). Consonant and vowel perception and production: early English-French bilinguals and English monolinguals. *Percept. Psychophys.* 46, 187–200. doi: 10.3758/BF03204982

MacLeod, A. N., and Stoel-Gammon, C. (2005). Are bilinguals different? What VOT tells us about simultaneous bilinguals. *J. Multicult. Commun. Disord.* 3, 118–127. doi: 10.1080/14769670500066313

MacWhinney, B. (2005). "A unified model of language acquisition," in *Handbook of Bilingualism: Psycholinguistic Approaches*, eds J. F. Kroll and A. M. B. de Groot (New York, NY: Oxford University Press), 49–67.

Marchman, V. A. (1993). Constraints on plasticity in a connectionist model of the English past tense. *J. Cogn. Neurosci.* 5, 215–234. doi: 10.1162/jocn.1993.5.2.215

May, S. (2014). *The Multilingual Turn: Implications for SLA, TESOL and Bilingual Education*. New York, NY: Routledge.

Mella, N., Fagot, D., and de Ribaupierre, A. (2016). Dispersion in cognitive functioning: age differences over the lifespan. *J. Clin. Exp. Neuropsychol.* 38, 111–126. doi: 10.1080/13803395.2015.1089979

Meulman, N., Wieling, M., Sprenger, S. A., Stowe, L. A., and Schmid, M. S. (2015). Age effects in L2 grammar processing as revealed by ERPs and how (not) to study them. *PLoS One* 10:e0143328. doi: 10.1371/journal.pone. 0143328

Montrul, S. (2016a). "Dominance and proficiency in early and late bilingualism," in *Language Dominance in Bilinguals: Issues of Measurement and Operationalization*, eds C. Silva-Corvalán and J. Treffers-Daller (Cambridge: Cambridge University Press), 15–35. doi: 10.1017/CBO9781107375345.002

Montrul, S. (2016b). "Age of onset of bilingualism effects and availability of input in first language attrition," in *Bilingualism Across the Lifespan: Factors Moderating Language Proficiency*, eds E. Nicoladis and S. Montanari (Washington, DC: American Psychological Association), 141–161.

Morgan-Short, K., Faretta-Stutenberg, M., Brill-Schuetz, K. A., Carpenter, H., and Wong, P. C. M. (2014). Declarative and procedural memory as individual differences in second language acquisition. *Bilingualism* 17, 56–72. doi: 10.1017/S1366728912000715

Moyer, A. (2014). Exceptional outcomes in L2 phonology: the critical factors of learner engagement and self-regulation. *Appl. Linguist.* 34, 418–440. doi: 10.1093/applin/amu012

Muñoz, C., and Singleton, D. (2011). A critical review of age-related research on L2 ultimate attainment. *Lang. Teach.* 44, 1–35. doi: 10.1017/S0261444810000327

Onnis, L., Chun, W. E., and Lou-Magnuson, M. (2018). Improved statistical learning abilities in adult bilinguals. *Bilingualism* 21, 427–433. doi: 10.1017/S1366728917000529

Ortega, L. (2009). *Understanding Second Language Acquisition*. London: Hodder.

Pallier, C. (2007). "Critical periods in language acquisition and language attrition," in *Language Attrition: Theoretical Perspectives*, eds B. Köpke, M. S. Schmid, M. Keijzer, and S. Dostert (Amsterdam: John Benjamins), 155–168. doi: 10.1075/sibil.33.11pal

Penfield, W., and Roberts, L. (1959). *Speech and Brain Mechanisms*. Princeton, NJ: Princeton University Press.

Petitto, L. A., Berens, M. S., Kovelman, I., Dubins, M. H., Jasinska, K., and Shalinsky, M. (2012). The "perceptual wedge hypothesis" as the basis for bilingual babies' phonetic processing advantage: new insights from fNIRS brain imaging. *Brain Lang.* 121, 130–143. doi: 10.1016/j.bandl.2011.05.003

Pinker, S. (1994). *The Language Instinct*. New York, NY: William Morrow. doi: 10.1037/e412952005-009

Pring, L. (2007). Savant talent. *Dev. Med. Child Neurol.* 47, 500–503. doi: 10.1017/S0012162205000976

Prull, M. W., Gabrieli, J. D. E., and Bunge, S. A. (1999). "Age-related changes in memory: a cognitive neuroscience perspective," in *The Handbook of Aging and Cognition*, eds F. I. M. Craik and T. A. Salthouse (Mahwah, NJ: Lawrence Erlbaum Associates), 91–153.

Pulvermüller, F., and Schumann, J. H. (1994). Neurobiological mechanisms of language acquisition. *Lang. Learn.* 44, 681–734. doi: 10.1111/j.1467-1770.1994. tb00635.x

Qureshi, M. A. (2016). A meta-analysis: age and second language grammar acquisition. *System* 60, 147–160. doi: 10.1016/j.system.2016. 06.001

Reichle, R. V., and Birdsong, D. (2014). Processing focus structure in L1 and L2 French: L2 proficiency effects on ERPs. *Stud. Second Lang. Acquis.* 36, 535–564. doi: 10.1017/S0272263113000594

Rossi, E., Díaz, M., Kroll, J. F., and Dussias, P. E. (2017). Late bilinguals are sensitive to unique aspects of second language processing: evidence from clitic pronouns word-order. *Front. Psychol.* 8:342. doi: 10.3389/fpsyg.2017.00342

Schmeißer, A., Hager, M., Gil, L. A., Jansen, V., Geveler, J., Eichler, N., et al. (2016). "Related but different: the two concepts of language dominance and language proficiency," in *Language Dominance in Bilinguals: Issues of Measurement and Operationalization*, eds C. Silva-Corvalán and J. Treffers-Daller (Cambridge: Cambridge University Press), 36–65.

Schwartz, A. I., and Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *J. Mem. Lang.* 55, 197–212. doi: 10.1016/j.jml.2006.03.004

Scovel, T. (1988). *A Time to Speak: A Psycholinguistic Inquiry into the Critical Period for Human Speech*. Rowley, MA: Newbury House.

Sebastián-Galles, N., and Díaz, B. (2012). First and second language speech perception: graded learning. *Lang. Learn.* 62(Suppl. 2), 131–147. doi: 10.1111/j. 1467-9922.2012.00709.x

Singleton, D. (2017). Language aptitude: desirable trait of acquirable attribute? *Stud. Second Lang. Learn. Teach.* 7, 89–103. doi: 10.14746/ssllt.2017. 7.1.5

Smith, N., Tsimpli, I., Morgan, G., and Woll, B. (2011). *The Signs of a Savant: Language Against the Odds*. Cambridge: Cambridge University Press.

Steinhauer, K. (2014). Event-related potentials (ERPs) in second language research: a brief introduction to the technique, a selected review, and an invitation to reconsider critical periods in L2. *Appl. Linguist.* 35, 393–417. doi: 10.1093/applin/amu028

Sundara, M., Polka, L., and Baum, S. (2006). Production of coronal stops by simultaneous bilingual adults. *Bilingualism* 9, 97–114. doi: 10.1017/S1366728905002403

Van Hell, J., and Abdollahi, F. (2017). "Individual variation in syntactic processing in the second language: behavioral and electrophysiological approaches," in *Developmental Perspectives in Written Language and Literacy*, eds E. Segers and P. van den Broek (Amsterdam: John Benjamins), 257–273. doi: 10.1075/z.206. 16van

Vanhove, J. (2013). The critical period hypothesis in second language acquisition: a statistical critique and a reanalysis. *PLoS One* 8:e69172. doi: 10.1371/journal. pone.0069172

Veríssimo, J., Heyer, V., Jacob, G., and Clahsen, H. (2017). Selective effects of age of acquisition on morphological priming: evidence for a sensitive period. *Lang. Acquis.* 1–12. doi: 10.1080/10489223.2017.1346104

Volkow, N. D., Gur, R. C., Wang, G. J., Fowler, J. S., Moberg, P. J., Ding, Y. S., et al. (1998). Association between decline in brain dopamine activity with age and cognitive and motor impairment in healthy individuals. *Am. J. Psychiatry* 155, 344–349. doi: 10.ll76/ajp.155.3.344

Werker, J. F., and Hensch, T. K. (2015). Critical periods in speech perception: new directions. *Annu. Rev. Psychol.* 66, 173–196. doi: 10.1146/annurev-psych-010814-015104

Wong, P. C. M., Morgan-Short, K., Ettlinger, M., and Zheng, J. (2012). Linking neurogenetics and individual differences in language learning: the dopamine hypothesis. *Cortex* 48, 1091–1102. doi: 10.1016/j.cortex.2012.03.017

Wong, P. C. M., Vuong, L. C., and Liu, K. (2017). Personalized learning: from neurogenetics of behaviors to designing optimal language training. *Neuropsychologia* 98, 192–200. doi: 10.1016/j.neuropsychologia.2016.10.002

Yow, W. Q., and Li, X. (2015). Balanced bilingualism and early age of second language acquisition as the underlying mechanisms of a bilingual executive control advantage: why variations in bilingual experiences matter. *Front. Psychol.* 6:164. doi: 10.3389/fpsyg.2015.00164

Uncovering the Mechanisms Responsible for Why Language Learning May Promote Healthy Cognitive Aging

Mark Antoniou* and Sarah M. Wright

The MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Sydney, NSW, Australia

One of the great challenges facing humankind in the 21st century is preserving healthy brain function in our aging population. Individuals over 60 are the fastest growing age group in the world, and by 2050, it is estimated that the number of people over the age of 60 will triple. The typical aging process involves cognitive decline related to brain atrophy, especially in frontal brain areas and regions that subserve declarative memory, loss of synaptic connections, and the emergence of neuropathological symptoms associated with dementia. The disease-state of this age-related cognitive decline is Alzheimer's disease and other dementias, which may cause older adults to lose their independence and rely on others to live safely, burdening family members and health care systems in the process. However, there are two lines of research that offer hope to those seeking to promote healthy cognitive aging. First, it has been observed that lifestyle variables such as cognitive leisure activities can moderate the risk of Alzheimer's disease, which has led to the development of plasticity-based interventions for older adults designed to protect against the adverse effects of cognitive decline. Second, there is evidence that lifelong bilingualism acts as a safeguard in preserving healthy brain function, possibly delaying the incidence of dementia by several years. In previous work, we have suggested that foreign language learning programs aimed at older populations are an optimal solution for building cognitive reserve because language learning engages an extensive brain network that is known to overlap with the regions negatively affected by the aging process. Here, we will outline potential future lines of research that may uncover the mechanism responsible for the emergence of language learning related brain advantages, such as language typology, bi- vs. multi-lingualism, age of acquisition, and the elements that are likely to result in the largest gains.

Keywords: bilingualism, language learning, cognitive aging, healthy aging, language typology

INTRODUCTION

One of the great challenges facing humankind in the twenty-first century is dealing with the problems associated with an aging population. Over-60-year-olds are the fastest growing age group on earth. By 2050, the number of people over 60 is set to triple, eclipsing 2 billion worldwide (Department of Economic and Social Affairs Population Division, 2007). As the number of older adults increases, so too will the demands and costs associated with an aging population, placing increasing pressure on families, health systems, economies, and governments. The typical aging

process is characterized by age-related decline in a number of cognitive subsystems (Park et al., 2002; Drachman, 2006). Certain brain structures are particularly affected by the aging process, such as frontal areas, the hippocampus, and the entorhinal cortex (Gómez-Isla et al., 1996; MacPherson et al., 2002; Bertoni-Freddari et al., 2003). Reduced function may be observed in working memory, declarative memory, as well as the interaction between declarative and procedural memory (Harrington and Haaland, 1992; Grady and Craik, 2000). The disease state of cognitive decline is Alzheimer's disease (and other dementias), characterized by a gradual progressive difficulty with learning and retaining new information. Pharmacological trials have had little success in slowing down the progression of Alzheimer's disease (Salloway et al., 2014). This has led to increasing calls to treat the disease proactively using behavioral stimulations, ideally before symptoms manifest (Selkoe, 2012).

Two promising lines of research have developed in parallel that offer some hope to combating age-related cognitive decline. On the one hand are studies demonstrating that environmental enrichment may result in positive brain changes. Studies of animals reared in standard vs. enriched enclosures have demonstrated the effects of environmental enrichment on the brain, namely denser dendritic connections resulting from stimulation (Volkmar and Greenough, 1972; Greenough et al., 1985). Such findings concerning environmental enrichment have been mirrored in investigations of lifestyle variables associated with healthy brain aging in humans: education, physical and mental stimulation, occupation, and leisure activities have all been linked to positive outcomes in cognitive aging (Kramer et al., 2004; Staff et al., 2004; Valenzuela and Sachdev, 2006; McDowell et al., 2007; Brayne et al., 2010; Foubert-Samier et al., 2012). These observations have led to the development of numerous plasticity-based interventions that aim to use environmental enrichment proactively by prescribing cognitively stimulating training regimens such as crossword puzzles (Verghese et al., 2003), math exercises (Kawashima et al., 2005), brain training (Ball et al., 2002), and computer-based interventions (Smith et al., 2009), and these cognitive improvements have been shown to persist over time (Mahncke, 2006). The resulting improvements have been observed in healthy adults, and encouragingly, also in those with mild cognitive impairment (Belleville et al., 2011), and even in those diagnosed with Alzheimer's disease (Bottino et al., 2005).

In parallel to the development and emergence of the cognitive training literature, evidence has been accumulating concerning the aging-related benefits of bilingualism. It was once thought that use of multiple languages (bilingualism) led to cognitive impairments (Goodenough, 1926). However, carefully conducted scientific studies later showed it to result in cognitive improvements (Peal and Lambert, 1962), an outcome since reinforced by a further 30 years of research. The evidence now suggests that experience with two languages confers a general 'bilingual advantage,' with improvement in executive function (Bialystok et al., 2004), metalinguistic awareness (Cummins, 1978), cognitive flexibility, creative thinking, and perhaps even several years' delay in the onset of dementia (Bialystok et al., 2007). Multilingualism is a better predictor of

cognitive ability than age, age at immigration, education, or sex (Kavé et al., 2008). These cognitive advantages that have been associated with bilingualism have neural correlates. For example, bilinguals demonstrate greater white matter integrity in old age compared to monolingual speakers (Luk et al., 2011). It has been suggested that this results in enhanced structural and functional connectivity that provides the neural basis for cognitive reserve. Bilingual older adults also show less steep cognitive decline than those who only speak one language (Bialystok, 2009). However, in recent years, the robustness of a bilingual advantage has been hotly debated—questioned by some who have failed to replicate it (Duñabeitia and Carreiras, 2015; Paap et al., 2015), but staunchly defended by its proponents (Bialystok et al., 2016). This debate itself highlights the absence of a detailed and systematic understanding of the factors that would underlie such an advantage. Interestingly, certain research laboratories consistently observe data patterns supporting a bilingual advantage, while other laboratories consistently find no advantage. It has even been suggested that bilingual advantages may reflect publication bias (de Bruin et al., 2015); but both the significant and non-significant findings are so systematic that it is much more likely that other factors (e.g., linguistic, experiential) are involved (Bialystok et al., 2015).

In the sections that follow, we will review the bilingual cognitive aging literature with a view to exploring the potential mechanisms responsible for the emergence of language learning related brain advantages and how these may be investigated prospectively in longitudinal language learning studies in adult learners.

BILINGUALISM AND EXECUTIVE FUNCTION

Many studies have reported that bilingualism yields advantages in executive function. Gold et al. (2013) found that bilingual older adults showed better task-switching performance than monolinguals in a color-shape task where participants categorize images by their color (blue or red) and shape (square and circle). Furthermore, fMRI scans taken during this task revealed decreased activation in the bilinguals' left lateral frontal cortex and cingulate cortex, an indication of more efficient executive functioning, when compared to a monolingual control group, and this difference was consistent across both the younger and older participants. Additionally, bilinguals outperformed monolinguals in episodic memory recall and letter fluency, but not the categorical fluency task (Ljungberg et al., 2013). Learners of French as a second language outperformed monolinguals on a grammaticality judgment task (ignoring conflict introduced through misleading semantic content) and a non-verbal visual search task (Janus et al., 2016). Older adult bilinguals, including those who acquired their second language in adulthood, exhibited improved cognitive function (general intelligence and reading) compared to monolinguals (Bak et al., 2014). When exposed to a non-verbal switching task, monolinguals showed activation in the right inferior frontal cortex and the anterior cingulate whereas bilinguals showed activation in the left inferior frontal

cortex and left striatum, both areas that underlie language control (Garbin et al., 2010). Older adult bilinguals processed distracting information more efficiently than their monolingual peers when completing the Flanker task (Ong et al., 2017). Aging bilinguals may also show less steep declines in executive function as they progress from healthy aging to mild cognitive impairment to probable Alzheimer's disease (Anderson et al., 2017). Collectively, these studies suggest that bilinguals show an advantage for non-linguistic cognitive abilities, particularly executive functions.

However, in recent years, a growing number of studies have raised questions regarding the robustness (and in some cases, the validity) of these bilingual advantage claims. This has led to attempts to understand the factors that give rise to the bilingual advantage in cognitive function and also explain its absence under certain conditions. Bilinguals' executive control abilities may be enhanced due to higher processing demands, and it has been argued that this may build cognitive reserve in the elderly (Costa and Sebastián-Gallés, 2014). Although the exact mechanisms are not agreed upon, it is generally thought that bilinguals' cognitive and brain reserves share the same mechanism as executive control processing (Grant et al., 2014). The complexity of the underlying cognitive processes may play a crucial role, with greater inhibitory demands resulting in greater benefit (Valian, 2015a,b). If correct, bilingualism may ultimately delay clinical Alzheimer's disease symptoms by protecting brain regions that subserve executive control (frontostriatal and frontoparietal) rather than those that subserve memory (medial temporal lobe) *per se* (Gold, 2015). Further, the potentially mediating effect of age of second language acquisition on executive functions is not well understood, and thus neither is its potential impact on the structure of the brain (Duñabeitia and Carreiras, 2015; Paap et al., 2015). Recently, Bialystok (2017) put forth an experience-dependent plasticity framework to evaluate the brain and cognitive modifications attributed to bilingualism. It was concluded that research broadly supports a relation between bilingualism and cognitive brain outcomes in infants and children, younger and older adults, and patients, however, behavioral studies with young adults, commonly fail to show these effects. This interpretation is consistent with findings in the executive function literature. Executive functions reach their peak in young adulthood (Park et al., 2002), and thus greater variability in executive functions, as measured by behavioral tasks, are more likely to be observed in older adulthood (when cognitive functions decline; Bialystok et al., 2008) or in childhood (when the foundations of cognitive processing are being established; Bayliss et al., 2003). Thus, it seems reasonable that bilingual advantages would be easier to detect either in early or later life.

BILINGUALISM AND COGNITIVE RESERVE

Cognitive reserve refers to the brain's resilience to neuropathological damage, resulting from experience-based neural changes associated with a physically and mentally stimulating lifestyle (Whalley et al., 2004). Stern (2012) proposes two possible mechanisms for cognitive reserve: neural reserve,

according to which differences in the resilience of already established networks, and neural compensation, according to which some individuals are better able to compensate for brain decline by using alternative networks. Evidence exists for both possibilities, and thus, the mechanisms responsible for cognitive reserve are a matter of ongoing research.

It is perhaps then unsurprising that the mechanism via which bilingualism improves the brain's resistance to neuropathology is not understood. Recent scholarly work has uncovered several potentially fruitful avenues concerning how bilingualism might build cognitive reserve focusing on the interactions between cognitive reserve and variables known to affect bilingualism (Calvo et al., 2016), as well as the brain networks that subserve memory (Grant et al., 2014), brain metabolic connectivity (Perani et al., 2017), and the presence of Alzheimer's disease biomarkers in cerebrospinal fluid (Estanga et al., 2017). Nevertheless, numerous studies present evidence suggesting that bilingualism results in brain changes in healthy subjects. Higher degrees of bilingualism have been linked to better lexical memory performance (Jafari et al., 2015). Bilinguals have higher white matter integrity than monolinguals in the corpus callosum extending to the superior and inferior longitudinal fasciculi, and also stronger anterior to posterior functional connectivity (Luk et al., 2011). Aging bilinguals outperformed monolinguals on the Flanker task, and had increased gray matter in the anterior cingulate cortex, whereas monolinguals showed decreased gray matter in the dorsolateral prefrontal cortex (Abutalebi et al., 2015b). Further, brain regions that support executive control significantly overlap with brain regions recruited for language control (Abutalebi and Green, 2016). The brain plasticity effects of lifelong bilingualism are thought to contribute to cognitive reserve and delay the onset of symptoms associated with dementia (Guzmán-Vélez and Tranel, 2015; Perani and Abutalebi, 2015). There is also evidence that bilingual brains are better able to accommodate anatomical and physiological brain changes and deterioration without exhibiting the expected increase in behavioral symptoms. Bilingual patients with Alzheimer's disease exhibited greater amounts of brain atrophy than monolingual patients (radial width of the temporal horn and the temporal horn ratio; Schweizer et al., 2012). Bilingual patients also showed substantially greater impairment of glucose uptake in frontotemporal and parietal regions (Brodmann areas 9, 47, 40, and 21) and in the left cerebellum relative to monolingual patients (Kowoll et al., 2016). This evidence supports the view that lifelong bilingualism may benefit the brain by making use of efficient or alternative neural networks in the event of age-related decline and that greater amounts of brain atrophy are required before the disease manifests, which may possibly delay the incidence of dementia.

DOES BILINGUALISM PROTECT AGAINST DEMENTIA?

The evidence for a protective effect of bilingualism on the incidence of dementia is considerable. Numerous studies have examined dementia incidence in hospital records and concluded

that bilingualism exerts a protective effect. The first such study by Bialystok et al. (2007) revealed that lifelong bilinguals showed a delay in the onset of symptoms of dementia by 4 years compared to monolinguals. Similarly, Craik et al. (2010) reported that bilingual patients had been diagnosed with Alzheimer's disease 4.3 years later and had reported the onset of symptoms 5.1 years later than the monolingual patients. Additionally, Woumans et al. (2015) found that bilingual patients had been diagnosed with Alzheimer's disease 4.8 years later and presented symptoms 4.6 years later than monolingual patients. Similarly, speakers of two or more languages had a delayed onset of Alzheimer's disease by up to 5 years and a protective effect was significant when speaking at least two to four languages (Freedman et al., 2014). Looking at specific dementia subtypes, bilingualism delayed the age at onset in the behavioral but not in the aphasic variants of Frontotemporal Dementia (Alladi et al., 2017), a finding consistent with the observation that bilingualism has positive effects on behavioral syndromes but not on language disorders. Indeed the effects of bilingualism on language functions are not always beneficial (e.g., smaller vocabulary size in a single language, slower lexical processing, reduced verbal fluency etc.). Further, a similar study by Alladi et al. (2016) comparing monolingual and bilingual stroke patients found that bilinguals had a significantly lower frequency of post-stroke dementia and mild cognitive impairment but the same frequency of post-stroke aphasia. Moreover, Atkinson (2016) reviewed nine papers and concluded that frequent use of two languages over a lifetime may be protective against dementia, and that inconsistencies arise due to study design or definitions of bilingualism. This evidence supports the protective effect of bilingualism against the symptoms of dementia (Bialystok et al., 2016), as well as the later onset of symptoms of mild cognitive impairment compared to monolinguals (Bialystok et al., 2014). Bilingual individuals diagnosed with single-domain amnesic mild cognitive impairment demonstrated a later age of diagnosis than did monolinguals (Ossher et al., 2013). Cerebral hypometabolism was more severe in the left hemisphere in bilinguals with Alzheimer's dementia compared to monolinguals, but nevertheless bilinguals outperformed monolinguals on memory tasks, suggesting that bilinguals are better able to compensate for the loss of brain structure and function (Perani et al., 2017). Furthermore, exposure to foreign language instruction during childhood and adolescence has been associated with lower risk of developing mild cognitive impairment in old age (Wilson et al., 2015). Bilingualism has been associated with delayed onset of dementia and is also observed in illiterate patients (Alladi et al., 2013). Taken together, this body of work suggests that bilingual experience delays the onset of neurodegenerative disease.

However, an increasing number of studies have failed to detect a bilingual advantage in dementia incidence. A cohort design with non-immigrant samples found no significant differences in the onset of dementia between mono- and bilingual subjects (Lawton et al., 2015). No significant association was found between non-native English speakers and the incidence of dementia or Alzheimer's disease (Sanders et al., 2012). In that study, non-native English speakers with at least 16 years of education had

a fourfold *increased risk* for dementia compared to those with less education, which is an unusual finding and inconsistent with past literature on the protective effect of education. Yeung et al. (2014) found no association between dementia diagnoses for bilinguals (English as a second language and bilingual English) and monolinguals. Zahodne et al. (2014) reported that adult learners of English had better memory and executive function than monolinguals, but that bilingualism was not associated with cognitive decline or dementia. Fuller-Thomson (2015) has claimed that the support for a bilingual advantage in dementia onset is questionable, and has attributed the current state of the literature to the file drawer problem, a bias against publishing non-significant findings from small studies with low to medium statistical power, a selection bias due to use of patients from a memory clinic, potential recall bias in caregivers' reporting of age of onset of dementia and confounding by immigration status. Indeed, Clare et al. (2016) did not observe any advantage for delay in Alzheimer's onset in Welsh-English bilinguals over English monolinguals (but see Bak, 2016 for a discussion of how this finding is conflated by the unusual situation of monolingual migration). A recent meta-analysis concluded that bilingualism offers no protection against cognitive decline (Mukadam et al., 2017), and that retrospective studies supporting the bilingual protective effect against dementia are marred by methodological confounds. Note, however, that this meta-analysis has already been criticized as misleading and incomplete (Woumans et al., 2017). In sum, these studies have led to questions regarding the robustness (or in some cases the validity) of the bilingual dementia advantage.

In order to resolve the debate, attempts have been made to understand the role of any potential mediating factors and experimental confounds. Gollan et al. (2011) claim that higher degrees of bilingualism are associated with increasingly later age of diagnosis and symptom onset, but this may be obscured by interactions between education and bilingualism, and a failure to obtain objective measures of bilingualism. Bak and Alladi (2014) highlight that although there exists support that bilingualism has a positive effect on cognition throughout the lifespan, common misconceptions concerning the nature of bilingualism persist, including that bilingualism is an unusual phenomenon, the holistic nature of bilingualism and its effects on cognition and bilingual diversity. Further, Fuller-Thomson's (2015) and Lawton et al.'s (2015) assertions that monolinguals and bilinguals do not differ in the onset of dementia have been criticized as overly simplistic. Bak and Alladi (2016) point out that it is necessary to study the effects of bilingualism separately from those of immigration and education, and to use data from both community-based approaches and memory clinics. Bak (2016) further highlights the importance of addressing confounding variables in bilingualism, aging and dementia research which include heterogeneity, migration, social factors, differences in general intelligence and the related issue of reverse causality.

The above literature review has demonstrated that bilingualism yields executive functioning advantages, and these may contribute to building cognitive reserve, which may ultimately delay the onset of dementia. The exact mechanisms are not agreed upon, and there exists counterevidence that limits

the generalisability of these claims. A possible fruitful avenue is the recent suggestion that sustained activation of noradrenergic signaling pathways associated with bilingualism could provide a possible mechanism linking current and previous results supporting a delayed onset of dementia in bilinguals (Bak and Robertson, 2017). The following sections of this article are devoted to proposing additional possible explanations and mechanisms that may provide parsimonious explanations for the seemingly conflicting findings currently in the literature.

AGE OF ACQUISITION

The majority of studies examining a bilingual advantage in cognitive aging have considered the effects of lifelong experience on cognitive function and decline. Consequently, very little attention has been paid to the age of acquisition of the second language. Age of acquisition of the second language positively correlates with cortical thickness in the left inferior frontal gyrus and a thinner cortex in the right inferior frontal gyrus (Klein et al., 2014). Encouragingly, there is evidence of a positive effect of language experience on individuals who acquired their languages later in life. Both early and late bilinguals were found to have more efficient executive networks than monolinguals. Proficient late bilinguals showed the greatest advantage in conflict resolution, whereas early bilinguals showed enhanced monitoring processes (Tao et al., 2011). Interestingly, Abutalebi et al. (2015a) found that age of acquisition did not correlate with gray matter volumes in the left or right inferior parietal lobules in aging bilinguals.

Age of acquisition is a complex variable in that it not only represents the level of input experienced by a learner, where early age of acquisition results in more years of exposure, but also potentially differing patterns of language use between speakers who acquired their second language in early or in later life. Such differences in language use may modulate the cognitive advantages associated with bilingualism. For instance, balanced bilinguals showed age-related decline in their inhibition abilities (as indexed by the Simon task), whereas dominant bilinguals showed no evidence of age-related decline (Goral et al., 2015). Further, when looking purely at amount of input, age of acquisition may need to be evaluated differently in older adulthood than it is for younger adults. For example, Tao et al. (2011) define early acquisition as occurring by an average of 4.0 years, and late acquisition as occurring by an average of 12.3 years. Although this difference in years of second language input might be marked for young adults, it is possible that this difference is negligible for those over the age of 65. Additionally, age of acquisition may result in executive control differences, not because of biological or maturational constraints on language learning, but because age of acquisition may be a proxy for a set of environmental differences that are necessarily associated with early vs. late second language learning (Tao et al., 2011). Indeed, those learning a language later due to migration will necessarily use their languages differently than someone learning a heritage language at an early age. Future longitudinal language training studies are needed to determine how age of acquisition modulates

any cognitive improvements resulting from language learning, and whether it truly is never too late to begin language learning.

NEUROIMAGING STUDIES OF LANGUAGE LEARNING IN ADULTS

A large neuroscientific literature has demonstrated that lifelong bilingualism alters the structure of the brain. Recent work has confirmed that brain changes may also be observed in healthy adults following relatively short periods of language training, and a picture is emerging concerning the brain changes that subserve dynamic uses of language (see **Table 1** for a summary of these findings). Interpreters who learned a foreign language intensively for 3 months showed increases in hippocampus volume and in cortical thickness in the left middle frontal gyrus, inferior frontal gyrus, and superior temporal gyrus, relative to controls; those with high proficiency showed structural malleability (right hippocampus and the left superior temporal gyrus) and struggling interpreters presented larger gray matter increases in the middle frontal gyrus (Mårtensson et al., 2012). Foreign language training in students increased white matter including pathways in the right hemisphere, and correlated with gain in second language ability not observed in controls (Hosoda et al., 2013). English natives who spent 5 months learning Swiss German showed structural changes in the left inferior frontal gyrus which correlated with increased second language proficiency (Stein et al., 2012). Moreover, structural changes in gray matter (inferior parietal cortex and left inferior frontal gyrus) and white matter (anterior corpus callosum) have been repeatedly linked to second language proficiency (Stein et al., 2014). Successful learners of a tonal language showed significant differences in language-related regions in the brain and a more coherent, integrated multi-path brain network compared to less successful learners, whereas monolinguals relied on different brain networks to process tonal and lexical information (Yang et al., 2015).

In sum, support has been found for second language experience-induced brain changes via increased gray matter density and white matter integrity in children, young adults, and the elderly; with such changes occurring rapidly following short-term language training. Further, these changes are sensitive to age, age of acquisition, proficiency or performance level, language-specific characteristics, and individual differences (Li et al., 2014).

THE ROLE OF LANGUAGE TYPOLOGY

One factor that has so far hardly played a role in the bilingual advantage debate concerns the languages that bilinguals use. A powerful factor might be the match or mismatch in typology (types of language structure, e.g., where verbs occur, use of affixes, etc.). Such structural features indeed influence learning of a foreign language (Cenoz, 2003; Antoniou et al., 2015), so they may also affect the likelihood of language-related advantages emerging, though this has not yet been examined.

TABLE 1 | Summary of empirical studies investigating brain changes in healthy adults due to language training.

Study	Sample	Training type	Training duration	Results
Hosoda et al., 2013	**24 language learners:** Japanese speaking students in basic level English courses (M_{age} = 20.1). **20 controls:** Japanese speaking students in basic level English courses. Age matched. Not participating in e-learning.	Computer-based e-learning.	60 idioms/words per week including spelling, meaning and pronunciation for 16 weeks.	Gray matter volume increased in right IFGop. White matter increased in right sub IFGop. Increased connectivity between IFGop caudate head pathway and dorsal pathway in RH.
Mårtensson et al., 2012	**14 conscript interpreters:** (M_{age} = 19.9), learning Arabic (n = 4), Dari (n = 8), or Russian (n = 2). No prior knowledge of these languages. **17 controls:** Medical and cognitive science students (M_{age} = 20.6).	Face-to-face classes and individual language studies.	Participants learned the language to fluency in 10 months. Scans conducted prior to start and after 3 months training. Participants learned 300–500 new words per week, and studied daily.	Increased cortical thickness in LH for dorsal MFG, IFG, STG, and right MFG and IFG. Greater hippocampal volume increase for interpreters, with proficiency related to right hippocampal volume and thickness of left STG. Changes in MIFG cortical thickness positively related to teacher ratings of student effort.
Stein et al., 2012	**10 English-speaking exchange students:** Learning German in Switzerland (M_{age} = 17.5).	Immersion and face-to-face classes.	5 months of second language learning including intensive 3-week course.	German proficiency related to increase in gray matter density for left IFG and ATL.
Yang et al., 2015	**21 English-speaking language learners:** No history of learning a tone language (M_{age} = 20.61). **13 passive controls:** No history of learning a tone language. Completed no language training (M_{age} = 20.8).	Computer-based.	Three 30 min training sessions per week for 6 weeks. Participants learned 48 Mandarin pseudoword picture pairs.	Increased activation of bilateral posterior MTG/AG for learners indicating that they treated tonal information as lexical. At completion of training, learners showed decreased activation for bilateral MEFG, MIFG, IFG, SMA, ACC, insula, and MTG relative to non-learners.

ACC, anterior cingulate cortex; AG, angular gyrus; ATL, anterior temporal lobe; IFG, inferior frontal gyrus; IFGop, inferior frontal gyrus pars opercularis; LH, left hemisphere; MEFG, medial frontal gyrus; MIFG, middle frontal gyrus; MTG, middle temporal gyrus; RH, right hemisphere; SMA, supplementary motor area; STG, superior temporal gyrus.

There are two possible, and contrasting, roles that language typology could play in determining whether cognitive advantages result from foreign language learning. First, typologically different languages might be more demanding to learn because they share few linguistic commonalities. It has recently been proposed that tasks that are more cognitively engaging will yield greater cognitive benefits (e.g., photographic training is superior to watching documentaries; Park et al., 2014), and for language learners, the benefits may be greatest when demands exceed their available cognitive resources (Schroeder and Marian, 2016). Second, typologically similar languages could lead to rapid learning because they share linguistic commonalities (and cognates), leading, after attainment of some proficiency, to competition between the languages exceeding that between distant languages (Weber and Cutler, 2004; Broersma and Cutler, 2011; Cutler, 2015). This in turn would require greater suppression, placing greater demands on the executive function system and its associated brain structures (prefrontal cortex; Stein et al., 2012, inferior parietal lobule; Mechelli et al., 2004, anterior cingulate; Abutalebi et al., 2012, basal ganglia; Zou et al., 2012, and putamen; Abutalebi et al., 2013). These alternatives predict opposite results regarding the appearance of cognitive benefit, but in both cases, the relationship between a learner's native and target languages would influence the demands on their cognitive resources.

We refer to the first possibility as the *processing complexity effect*, according to which greater cognitive improvements will ensue from learning typologically differing languages, as these require more effort to learn and existing native-language knowledge cannot be relied on. The alternative is the *interference inhibition effect*, according to which greater cognitive improvements will ensue from learning typologically similar languages, because similar languages interfere more, increasing demands on executive control systems in the brain.

As noted above, studies to date have typically reported advantages for individuals using multiple languages for many years. But rigorous investigation of how language learning affects cognitive function must: (a) measure the cognitive abilities of interest *prior to*, as well as post, language learning, and (b) experimentally manipulate who learns what language. Neither can be done with individuals who are already bilingual. Systematic investigation of the effects outlined above will require longitudinal experiments in which participants will be cognitively assessed both prior to and after completing language training. This allows (1) control of extraneous variables (e.g., education, Gollan et al., 2011;

socioeconomic status, Calvo and Bialystok, 2014; quality and quantity of language input, Sorace and Serratrice, 2009), (2) assignment of participants to target language, and (3) experimental manipulation of the relationship between the target language and the learner's native language (i.e., typological similarity).

Although these possibilities have not yet been tested systematically, there is some evidence that typological distance will result in reliable brain differences. For example, Abutalebi et al. (2015a) observed greater gray matter volumes in the brains of aging Chinese bilinguals relative to monolinguals, specifically in the left and right inferior parietal lobules. Importantly, when comparing Cantonese-English and Cantonese-Mandarin bilinguals, both groups showed greater gray matter volumes for the right inferior parietal lobule, but only Cantonese-Mandarin bilinguals showed greater gray matter volumes for the left inferior parietal lobule. Although preliminary, this observation is consistent with our hypothesized interference inhibition effect, suggesting that two similar languages result in greater competition and place greater demands on the executive control system, requiring more inhibition to avoid language interference. This may result in brain differences that are more prominent for speakers of typologically similar languages. This interference inhibition effect could potentially go some way to providing a parsimonious explanation for conflicting findings in the literature.

THE COGNITIVE BENEFITS OF ADDITIONAL LANGUAGE LEARNING: BILINGUALS vs. MULTILINGUALS

Experience with multiple languages is thought to yield cognitive advantages that promote healthy cognitive aging, although the mechanisms responsible are not fully understood. If our proposed interference inhibition effect is correct, then conditions that increase interference will yield the greatest benefits. Therefore, knowledge of a greater number of languages would likely increase competition and the required inhibitory control and thus yield a greater cognitive benefit. However, we note that presently it is not clear if experience with a greater number of languages results in an additive benefit. There is some evidence suggesting that this may indeed be the case. Aging bilinguals and multilinguals maintain higher levels of cognitive functioning than monolinguals, irrespective of immigration and education levels (Kavé et al., 2008). Chertkow et al. (2010) found evidence for a later age of onset of Alzheimer's disease symptoms in multilinguals as compared with monolinguals, whereas a limited effect was found between bilinguals compared to monolinguals. Participants who practiced more than two languages presented a lower risk of cognitive impairment without dementia, compared to bilinguals. Progressing from two to three languages was associated with a sevenfold protection against cognitive impairment without dementia (Perquin et al., 2013). There is some evidence that this multilingual benefit is mediated by age. Older trilingual adults showed larger advantages on cognitive reserve than bilinguals. However, younger trilingual adults and children showed the same advantages as bilinguals on inhibitory control measures. Trilingual infants and toddlers performed worse than bilinguals on memory generalization tasks (Schroeder and Marian, 2016). In sum, it is not clear if multilingualism brings about greater cognitive benefits than bilingualism, although the present evidence suggests that it is likely to emerge under certain circumstances.

LANGUAGE LEARNING STUDIES WITH OLDER ADULTS

We have established that there is evidence to suggest that lifelong bilingualism may enhance executive functions, contribute to cognitive reserve, and possibly protect against Alzheimer's disease. However, whether language learning initiated in older adulthood could yield cognitive improvements remains an open research question. Previous research has shown that both healthy older adults and those at risk of neural dysfunction have demonstrated positive brain changes to training (Valenzuela et al., 2003). This indicates that the benefits of mental stimulation are not limited to younger adults, but that even the aging brain retains its neuroplasticity, and thus training-related benefits may still be observed in older participants. Given that language learning engages an extensive neural network (Rodríguez-Fornells et al., 2009) that overlaps with the network affected by age-related cognitive decline (Raz, 2000), a tantalizing possibility is that language learning may promote healthy brain aging in older adults (see Antoniou et al., 2013 for a full review). Although there is currently very little research in this area, there are positive signs that this may well be the case (**Table 2**). For example, Bak et al. (2016) found language learning advantages for task switching using the elevator task with reversal. In this task, participants listen to a sequence of three different tones: low, mid, and high. Participants are required to count the mid tones, add one for the high tones, and subtract one for the low tones. The advantage demonstrated on this task was found after only 1 week of intensive language training, composed of 14 h of formal classes in Scottish Gaelic. Participants were also offered Gaelic entertainment in the evenings including concerts, films and conversation circles. In addition, it was found that those individuals who continued practicing Gaelic for at least 5 h per week following the cessation of the course retained their improvement at the 9 month follow-up. Finally, and perhaps most importantly, improved attentional switching was observed across all age groups, ranging from 18 to 78 years of age, indicating that just 1 week of foreign language training can provide some cognitive benefit even for older learners. In contrast, Ramos et al. (2017) did not observe an improvement in non-verbal task switching ability in older Spanish monolinguals who participated in 8 months of Basque language classes for 5.5 h per week. However, it is not clear why these authors elected to examine task switching as an outcome measure, rather than for example inhibitory control, which would be expected to show some training-related changes following several months of language learning. The domain-general brain circuitry that subserves task switching will not necessarily be affected by

TABLE 2 | Summary of studies investigating language learning benefits in older adults.

Study	Sample	Training type	Training duration	Results
Bak et al., 2016	**33 language learners:** intensive Scottish Gaelic course. **16 active controls:** Students of non-language courses. **18 passive controls:** Recruited from Psychology volunteer pool.	Face-to-face intensive course.	14 h over a 1 week period.	Language learners showed advantage for attentional switching using the Elevator Task with Reversal at end of course. Improved switching maintained at 9 months follow-up for those practicing Gaelic for at least 5 h/week after the cessation of the course for all age groups (18–78 years).
Ramos et al., 2017	**26 language learners:** Spanish monolinguals learning Basque. Aged between 60 and 80 (M_{age} = 67.42). **17 passive controls:** Spanish monolinguals. Aged between 60 and 78 (M_{age} = 69.18).	Face-to-face, small group lessons (maximum 10 participants).	5.5 h (3 days) per week for 8 months.	No decrease in switching cost using the Color-Shape task.
Ware et al., 2017	**14 language learners:** French speaking older adults (M_{age} = 75 years) with varying levels of English language knowledge.	Face-to-face group lessons with integrated technology (laptop/tablet).	2 h per week for 4 months. Participants were encouraged to practice using their laptops/tablets outside of sessions.	No difference in cognitive level, loneliness or social isolation. Demonstrated that a technology-based language learning intervention of this duration and intensity is viable for use with this population.

language learning *per se*, but rather will depend critically on a bilingual's pattern of language use. For instance, constant switching between languages (as in the case of codeswitching), or the need to constantly monitor the environment for both languages would be expected to yield improvements in task switching more broadly (Green and Abutalebi, 2013). There are three key differences between the studies conducted by Bak et al. (2016) and Ramos et al. (2017) that may give way to differing outcomes in terms of task switching improvement. The first is the intensity of the initial training. Participants in the Ramos et al. (2017) study had less intense training in the initial week. That is, they participated in three sessions totalling 5.5 h. In comparison, the participants in the Bak et al. (2016) study completed approximately 14 h of language classes, and additional Gaelic language activities. Beyond the initial week, participants in both studies completed at least 5 h of practice in their respective languages. The second difference is that each study measured switching with a different task. Ramos et al. (2017) used the Color-Shape Task, a task commonly used to measure shifting between mental sets, whereas Bak et al. (2016) used the Elevator Task with Reversal, a measure of attentional switching from the Test of Everyday Attention. This can be problematic as these two tasks stem from different theoretical perspectives (the former from working memory, and the latter from attention research) it is not known whether these two tasks measure comparable constructs (see Mackie et al., 2013 for a review on defining cognitive control and attentional functions). The third difference between these studies that may give way to improvement differences, is the context of subsequent language use. While participants in the Ramos et al. (2017) study continued formal classes for 5.5 h per week, those in the Bak et al. (2016) no longer continued their intensive language training. However, it was only those that continued to use Gaelic for at least 5 h per week that improved from their baseline switching performance. Given that language switching is expected to provide improvements

to switching more broadly, it is likely that formal classes such as those used in the Ramos et al. (2017) study provide less opportunity for codeswitching, compared to the Gaelic learners who were practicing Gaelic in their everyday lives.

Finally, a recent second language training study aimed to determine whether an English learning program implemented with French-speaking seniors would improve cognition, as well as subjective levels of loneliness and social isolation. Scores on these measures did not improve significantly, perhaps due to the small sample size or short study duration, including the length of the language learning sessions themselves. However, the study did demonstrate that a 2-h per week, technology-based language learning intervention is feasible for seniors to participate in (Ware et al., 2017). Given that Bak et al. (2016) determined that 5 h per week is the minimum level of language use required for cognitive advantages to arise, future research in this area needs to determine whether this also extends to technology-based interventions. Additionally, further research investigating our proposed processing complexity and interference inhibition effects will assist researchers in determining if typological similarity can be used to maximize language training for aging populations. Whether language learning can yield cognitive improvements in older adults, and if so, under what specific conditions, remain open research questions. Answers to these questions are being pursued by research laboratories around the world.

LANGUAGE USE IN INDIVIDUALS WITH ALZHEIMER'S DISEASE

One final research question concerns the potential role of language learning in individuals with mild cognitive impairment or Alzheimer's disease. Studies with Alzheimer's patients often suffer from design inconsistencies and small sample sizes.

However, a picture is starting to emerge regarding bilingual language use in Alzheimer's disease. Bilinguals diagnosed with Alzheimer's disease may exhibit cognitive impairment and lapses in attention, decreased language control ability and increased unwanted code-switching (Friedland and Miller, 1999). Bilingual individuals with Alzheimer's disease show linguistic decrements in both their dominant and non-dominant languages (Stilwell et al., 2016). English dominant bilinguals with Alzheimer's disease were more likely to name pictures in the non-dominant language than controls; and Spanish-dominant bilinguals with Alzheimer's disease were equally likely to name pictures in their non-dominant language than controls (Gollan et al., 2011). A case study of two bilingual patients presented early symptoms of dementia after regressing to their primary language (McMurtray et al., 2009). Bilinguals with mild to moderate dementia had impaired retrieval of their first language (Frisian) and L2 (Dutch) naming ability, with a significant effect of age of acquisition. Earlier acquired words were better preserved and retrieved. Qualitatively, inappropriate code switching occurred within the Frisian test setting (Veenstra et al., 2014). These studies provide a glimpse of the effects of Alzheimer's disease on bilingual language use. Whether language training benefits Alzheimer's disease patients warrants future investigation.

CONCLUSION

In this review, we have outlined the benefits of bilingualism on executive functioning and how this may increase cognitive reserve in older adults. Additionally, we have discussed how foreign language learning programs may potentially promote healthy aging and protect against cognitive decline including Alzheimer's disease, as a result of the overlap between the brain networks involved in language learning and those that decline in older age. It is proposed that future research in this area should aim to uncover the mechanisms responsible

for language learning related brain advantages, and determine how language learning can be optimized to reap the maximum cognitive gains. Specifically, to achieve these aims, future research should determine the role that language typology plays in promoting healthy cognitive aging by systematically manipulating typological similarity in foreign language learning studies. In doing so, language learning programs can be customized to provide maximal cognitive advantage in line with either the processing complexity or interference inhibition effects. We also suggest that more rigorous investigation in this field could be achieved by measuring cognitive abilities prior to, as well as post language learning. Further, we have discussed the potential advantages of bilingualism vs. multilingualism, and suggest that studies that compare cognitive advantages of language learning between monolinguals learning a second language, and bilinguals learning a third, could reveal whether learning additional languages provides an additive effect. Finally, future research needs to determine the optimum language learning conditions that will provide maximum cognitive benefits in older populations. The findings from these lines of research would provide convincing evidence as to whether language learning might promote healthy cognitive aging in older adulthood and, if so, provide guidelines for how these programs should be developed to provide the greatest cognitive advantage.

AUTHOR CONTRIBUTIONS

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

ACKNOWLEDGMENT

REFERENCES

Abutalebi, J., Canini, M., Della Rosa, P. A., Green, D. W., and Weekes, B. S. (2015a). The neuroprotective effects of bilingualism upon the inferior parietal lobule: a structural neuroimaging study in aging Chinese bilinguals. *J. Neurolinguistics* 33, 3–13. doi: 10.1016/j.jneuroling.2014.09.008

Abutalebi, J., Della Rosa, P. A., Castro Gonzaga, A. K., Keim, R., Costa, A., and Perani, D. (2013). The role of the left putamen in multilingual language production. *Brain Lang.* 125, 307–315. doi: 10.1016/j.bandl.2012.03.009

Abutalebi, J., Della Rosa, P. A., Green, D. W., Hernandez, M., Scifo, P., Keim, R., et al. (2012). Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cereb. Cortex* 22, 2076–2086. doi: 10.1093/cercor/bhr287

Abutalebi, J., and Green, D. W. (2016). Neuroimaging of language control in bilinguals: neural adaptation and reserve. *Bilingualism* 19, 1–10. doi: 10.1017/S1366728916000225

Abutalebi, J., Guidi, L., Borsa, V., Canini, M., Della Rosa, P. A., Parris, B. A., et al. (2015b). Bilingualism provides a neural reserve for aging populations. *Neuropsychologia* 69, 201–210. doi: 10.1016/j.neuropsychologia.2015.01.040

Alladi, S., Bak, T. H., Duggirala, V., Surampudi, B., Shailaja, M., Shukla, A. K., et al. (2013). Bilingualism delays age at onset of dementia, independent of education and immigration status. *Neurology* 81, 1938–1944. doi: 10.1212/01.wnl.0000436620.33155.a4

Alladi, S., Bak, T. H., Mekala, S., Rajan, A., Chaudhuri, J. R., Mioshi, E., et al. (2016). Impact of bilingualism on cognitive outcome after stroke. *Stroke* 47, 258–261. doi: 10.1161/STROKEAHA.115.010418

Alladi, S., Bak, T. H., Shailaja, M., Gollahalli, D., Rajan, A., Surampudi, B., et al. (2017). Bilingualism delays the onset of behavioral but not aphasic forms of frontotemporal dementia. *Neuropsychologia* 99, 207–212. doi: 10.1016/j.neuropsychologia.2017.03.021

Anderson, J. A. E., Saleemi, S., and Bialystok, E. (2017). Neuropsychological assessments of cognitive aging in monolingual and bilingual older adults. *J. Neurolinguistics* 43, 17–27. doi: 10.1016/j.jneuroling.2016.08.001

Antoniou, M., Gunasekera, G. M., and Wong, P. C. M. (2013). Foreign language training as cognitive therapy for age-related cognitive decline: a hypothesis for future research. *Neurosci. Biobehav. Rev.* 37, 2689–2698. doi: 10.1016/j.neubiorev.2013.09.004

Antoniou, M., Liang, E., Ettlinger, M., and Wong, P. C. M. (2015). The bilingual advantage in phonetic learning. *Bilingualism* 18, 683–695. doi: 10.1017/s1366728914000777

Atkinson, A. L. (2016). Does bilingualism delay the development of dementia? *J. Eur. Psychol. Stud.* 7, 43–50. doi: 10.5334/jeps.375

Bak, T., and Alladi, S. (2014). Can being bilingual affect the onset of dementia? *Future Neurol.* 9, 1–6. doi: 10.2217/FNL.14.8

Bak, T. H. (2016). The impact of bilingualism on cognitive aging and dementia: finding a path through a forest of confounding variables. *Linguist. Approaches Biling.* 6, 205–226. doi: 10.1075/lab.15002.bak

Bak, T. H., and Alladi, S. (2016). Bilingualism, dementia and the tale of many variables: why we need to move beyond the Western World. Commentary on Lawton et al. (2015) and Fuller-Thomson (2015). *Cortex* 74, 315–317. doi: 10.1016/j.cortex.2015.04.025

Bak, T. H., Long, M. R., Vega-Mendoza, M., and Sorace, A. (2016). Novelty, challenge, and practice: the impact of intensive language learning on attentional functions. *PLOS ONE* 11:e0153485. doi: 10.1371/journal.pone.0153485

Bak, T. H., Nissan, J. J., Allerhand, M. M., and Deary, I. J. (2014). Does bilingualism influence cognitive aging? *Ann. Neurol.* 75, 959–963. doi: 10.1002/ana.24158

Bak, T. H., and Robertson, I. (2017). Biology enters the scene—a new perspective on bilingualism, cognition, and dementia. *Neurobiol. Aging* 50, iii–iv. doi: 10.1016/j.neurobiolaging.2016.10.020

Ball, K., Berch, D. B., Helmers, K. F., Jobe, J. B., Leveck, M. D., Marsiske, M., et al. (2002). Effects of cognitive training interventions with older adults: a randomized controlled trial. *J. Am. Med. Assoc.* 288, 2271–2281. doi: 10.1001/jama.288.18.2271

Bayliss, D. M., Jarrold, C., Gunn, D. M., and Baddeley, A. D. (2003). The complexities of complex span: explaining individual differences in working memory in children and adults. *J. Exp. Psychol. Gen.* 132, 71–92. doi: 10.1037/0096-3445.132.1.71

Belleville, S., Clément, F., Mellah, S., Gilbert, B., Fontaine, F., and Gauthier, S. (2011). Training-related brain plasticity in subjects at risk of developing Alzheimer's disease. *Brain* 134, 1623–1634. doi: 10.1093/brain/awr037

Bertoni-Freddari, C., Fattoretti, P., Solazzi, M., Giorgetti, B., Di Stefano, G., Casoli, T., et al. (2003). Neuronal death versus synaptic pathology in Alzheimer's disease. *Ann. N. Y. Acad. Sci.* 1010, 635–638. doi: 10.1196/annals.1229.116

Bialystok, E. (2009). Bilingualism: the good, the bad, and the indifferent. *Bilingualism* 12, 3–11. doi: 10.1017/S1366728908003477

Bialystok, E. (2017). The bilingual adaptation: how minds accommodate experience. *Psychol. Bull.* 143, 233–262. doi: 10.1037/bul0000099

Bialystok, E., Abutalebi, J., Bak, T. H., Burke, D. M., and Kroll, J. F. (2016). Aging in two languages: implications for public health. *Ageing Res. Rev.* 27, 56–60. doi: 10.1016/j.arr.2016.03.003

Bialystok, E., Craik, F. I. M., Binns, M. A., Ossher, L., and Freedman, M. (2014). Effects of bilingualism on the age of onset and progression of MCI and AD: evidence from executive function tests. *Neuropsychology* 28, 290–304. doi: 10.1037/neu0000023

Bialystok, E., Craik, F. I. M., and Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia* 45, 459–464. doi: 10.1016/j.neuropsychologia.2006.10.009

Bialystok, E., Craik, F. I. M., Klein, R. M., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19, 290–303. doi: 10.1037/0882-7974.19.2.290

Bialystok, E., Craik, F. I. M., and Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 859–873. doi: 10.1037/0278-7393.34.4.859

Bialystok, E., Kroll, J. F., Green, D. W., Macwhinney, B., and Craik, F. I. M. (2015). Publication bias and the validity of evidence: What's the connection? *Psychol. Sci.* 26, 944–946. doi: 10.1177/0956797615573759

Bottino, C. M., Carvalho, I. A., Alvarez, A. M., Avila, R., Zukauskas, P. R., Bustamante, S. E., et al. (2005). Cognitive rehabilitation combined with drug treatment in Alzheimer's disease patients: a pilot study. *Clin. Rehabil.* 19, 861–869. doi: 10.1191/0269215505cr911oa

Brayne, C., Ince, P. G., Keage, H. A. D., McKeith, I. G., Matthews, F. E., Polvikoski, T., et al. (2010). Education, the brain and dementia: neuroprotection or compensation? *Brain* 133, 2210–2216. doi: 10.1093/brain/awq185

Broersma, M., and Cutler, A. (2011). Competition dynamics of second-language listening. *Q. J. Exp. Psychol.* 64, 74–95. doi: 10.1080/17470218.2010.499174

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

Calvo, N., García, A. M., Manoiloff, L., and Ibáñez, A. (2016). Bilingualism and cognitive reserve: a critical overview and a plea for methodological innovations. *Front. Aging Neurosci.* 7:249. doi: 10.3389/fnagi.2015.00249

Cenoz, J. (2003). The additive effect of bilingualism on third language acquisition: a review. *Int. J. Biling.* 7, 71–87. doi: 10.1177/13670069030070010501

Chertkow, H., Whitehead, V., Phillips, N., Wolfson, C., Atherton, J., and Bergman, H. (2010). Multilingualism (but not always bilingualism) delays the onset of Alzheimer disease: evidence from a bilingual community. *Alzheimer Dis. Assoc. Disord.* 24, 118–125. doi: 10.1097/WAD.0b013e3181ca1221

Clare, L., Whitaker, C. J., Craik, F. I. M., Bialystok, E., Martyr, A., Martin-Forbes, P. A., et al. (2016). Bilingualism, executive control, and age at diagnosis among people with early-stage Alzheimer's disease in Wales. *J. Neuropsychol.* 10, 163–185. doi: 10.1111/jnp.12061

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

Craik, F. I. M., Bialystok, E., and Freedman, M. (2010). Delaying the onset of Alzheimer disease. *Neurology* 75, 1726–1729. doi: 10.1212/WNL.0b013e3181fc2a1c

Cummins, J. (1978). Bilingualism and the development of metalinguistic awareness. *J. Cross Cult. Psychol.* 9, 131–149. doi: 10.1177/002202217892001

Cutler, A. (2015). Representation of second language phonology. *Appl. Psycholinguist.* 36, 115–128. doi: 10.1017/S0142716414000459

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

Department of Economic and Social Affairs Population Division (2007). *World Population Ageing*. New York, NY: United Nations Publications.

Drachman, D. A. (2006). Aging of the brain and Alzheimer's disease. *Neurology* 67, 1340–1352. doi: 10.1212/01.wnl.0000240127.89601.83

Duñabeitia, J. A., and Carreiras, M. (2015). The bilingual advantage: acta est fabula? *Cortex* 73, 371–372. doi: 10.1016/j.cortex.2015.06.009

Estanga, A., Ecay-Torres, M., Ibáñez, A., Izagirre, A., Villanua, J., Garcia-Sebastian, M., et al. (2017). Beneficial effect of bilingualism on Alzheimer's disease CSF biomarkers and cognition. *Neurobiol. Aging* 50, 144–151. doi: 10.1016/j.neurobiolaging.2016.10.013

Foubert-Samier, A., Catheline, G., Amieva, H., Dilharreguy, B., Helmer, C., Allard, M., et al. (2012). Education, occupation, leisure activities, and brain reserve: a population-based study. *Neurobiol. Aging* 33, 423.e15–423.e25. doi: 10.1016/j.neurobiolaging.2010.09.023

Freedman, M., Alladi, S., Chertkow, H., Bialystok, E., Craik, F. I. M., Phillips, N. A., et al. (2014). Delaying onset of dementia: are two languages enough? *Behav. Neurol.* 2014:808137. doi: 10.1155/2014/808137

Friedland, D., and Miller, N. (1999). Language mixing in bilingual speakers with Alzheimer's dementia: a conversation analysis approach. *Aphasiology* 13, 427–444. doi: 10.1080/026870399402163

Fuller-Thomson, E. (2015). Emerging evidence contradicts the hypothesis that bilingualism delays dementia onset. A Commentary on "age of dementia diagnosis in community dwelling bilingual and monolingual Hispanic Americans" by Lawton et al., 2015. *Cortex* 66, 170–172. doi: 10.1016/j.cortex.2015.01.024

Garbin, G., Sanjuan, A., Forn, C., Bustamante, J. C., Rodriguez-Pujadas, A., Belloch, V., et al. (2010). Bridging language and attention: brain basis of the impact of bilingualism on cognitive control. *Neuroimage* 53, 1272–1278. doi: 10.1016/j.neuroimage.2010.05.078

Gold, B. T. (2015). Lifelong bilingualism and neural reserve against Alzheimer's disease: a review of findings and potential mechanisms. *Behav. Brain Res.* 281, 9–15. doi: 10.1016/j.bbr.2014.12.006

Gold, B. T., Kim, C., Johnson, N. F., Kryscio, R. J., and Smith, C. D. (2013). Lifelong bilingualism maintains neural efficiency for cognitive control in aging. *J. Neurosci.* 33, 387–396. doi: 10.1523/JNEUROSCI.3837-12.2013

Gollan, T. H., Salmon, D. P., Montoya, R. I., and Galasko, D. R. (2011). Degree of bilingualism predicts age of diagnosis of Alzheimer's disease in low-education but not in highly educated Hispanics. *Neuropsychologia* 49, 3826–3830. doi: 10.1016/j.neuropsychologia.2011.09.041

Gómez-Isla, T., Price, J. L., McKeel, D. W., Morris, J. C., Growdon, J. H., and Hyman, B. T. (1996). Profound loss of layer II entorhinal cortex neurons occurs in very mild Alzheimer's disease. *J. Neurosci.* 16, 4491–4500.

Goodenough, F. L. (1926). Racial differences in the intelligence of school children. *J. Exp. Psychol.* 9, 388–397. doi: 10.1037/h0073325

Goral, M., Campanelli, L., and Spiro, A. (2015). Language dominance and inhibition abilities in bilingual older adults. *Bilingualism* 18, 79–89. doi: 10.1017/S1366728913000126

Grady, C. L., and Craik, F. I. M. (2000). Changes in memory processing with age. *Curr. Opin. Neurol.* 10, 224–231. doi: 10.1016/S0959-4388(00)00073-8

Grant, A., Dennis, N. A., and Li, P. (2014). Cognitive control, cognitive reserve, and memory in the aging bilingual brain. *Front. Psychol.* 5:1401. doi: 10.3389/fpsyg.2014.01401

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

Greenough, W. T., Hwang, H. M., and Gorman, C. (1985). Evidence for active synapse formation or altered postsynaptic metabolism in visual cortex of rats reared in complex environments. *Proc. Natl. Acad. Sci. U.S.A.* 82, 4549–4552. doi: 10.1073/pnas.82.13.4549

Guzmán-Vélez, E., and Tranel, D. (2015). Does bilingualism contribute to cognitive reserve? Cognitive and neural perspectives. *Neuropsychology* 29, 139–150. doi: 10.1037/neu0000105

Harrington, D. L., and Haaland, K. Y. (1992). Skill learning in the elderly: diminished implicit and explicit memory for a motor sequence. *Psychol. Aging* 7, 425–434. doi: 10.1037/0882-7974.7.3.425

Hosoda, C., Tanaka, K., Nariai, T., Honda, M., and Hanakawa, T. (2013). Dynamic neural network reorganization associated with second language vocabulary acquisition: a multimodal imaging study. *J. Neurosci.* 33, 13663–13672. doi: 10.1523/JNEUROSCI.0410-13.2013

Jafari, Z., Esmaili, M., Toufan, R., and Aghamollaei, M. (2015). Bilingual proficiency and cognitive reserve in Persian-English bilingual older adults. *Aging Clin. Exp. Res.* 27, 351–357. doi: 10.1007/s40520-014-0288-x

Janus, M., Lee, Y., Moreno, S., and Bialystok, E. (2016). Effects of short-term music and second-language training on executive control. *J. Exp. Child Psychol.* 144, 84–97. doi: 10.1016/j.jecp.2015.11.009

Kavé, G., Eyal, N., Shorek, A., and Cohen-Mansfield, J. (2008). Multilingualism and cognitive state in the oldest old. *Psychol. Aging* 23, 70–78. doi: 10.1037/0882-7974.23.1.70

Kawashima, R., Okita, K., Yamazaki, R., Tajima, N., Yoshida, H., Taira, M., et al. (2005). Reading aloud and arithmetic calculation improve frontal function of people with dementia. *J. Gerontol. A Biol. Sci. Med. Sci.* 60, 380–384. doi: 10.1093/gerona/60.3.380

Klein, D., Mok, K., Chen, J. K., and Watkins, K. E. (2014). Age of language learning shapes brain structure: a cortical thickness study of bilingual and monolingual individuals. *Brain Lang.* 131, 20–24. doi: 10.1016/j.bandl.2013.05.014

Kowoll, M. E., Degen, C., Gorenc, L., Küntzelmann, A., Fellhauer, I., Giesel, F., et al. (2016). Bilingualism as a contributor to cognitive reserve? Evidence from cerebral glucose metabolism in mild cognitive impairment and Alzheimer's disease. *Front. Psychiatry* 7:62. doi: 10.3389/fpsyt.2016.00062

Kramer, A. F., Bherer, L., Colcombe, S. J., Dong, W., and Greenough, W. T. (2004). Environmental influences on cognitive and brain plasticity during aging. *J. Gerontol. A Biol. Sci. Med. Sci.* 59, 940–957. doi: 10.1093/gerona/59.9.M940

Lawton, D. M., Gasquoine, P. G., and Weimer, A. A. (2015). Age of dementia diagnosis in community dwelling bilingual and monolingual Hispanic Americans. *Cortex* 66, 141–145. doi: 10.1016/j.cortex.2014.11.017

Li, P., Legault, J., and Litcofsky, K. A. (2014). Neuroplasticity as a function of second language learning: anatomical changes in the human brain. *Cortex* 58, 301–324. doi: 10.1016/j.cortex.2014.05.001

Ljungberg, J. K., Hansson, P., Andrés, P., Josefsson, M., and Nilsson, L.-G. (2013). A longitudinal study of memory advantages in bilinguals. *PLOS ONE* 8:e73029. doi: 10.1371/journal.pone.0073029

Luk, G., Bialystok, E., Craik, F. I. M., and Grady, C. L. (2011). Lifelong bilingualism maintains white matter integrity in older adults. *J. Neurosci.* 31, 16808–16813. doi: 10.1523/JNEUROSCI.4563-11.2011

Mackie, M.-A., Van Dam, N. T., and Fan, J. (2013). Cognitive control and attentional functions. *Brain Cogn.* 82, 301–312. doi: 10.1016/j.bandc.2013.05.004

MacPherson, S. E., Phillips, L. H., and Della Sala, S. (2002). Age, executive function, and social decision making: a dorsolateral prefrontal theory of cognitive aging. *Psychol. Aging* 17, 598–609. doi: 10.1037//0882-7974.17.4.598

Mahncke, H. W. (2006). Memory enhancement in healthy older adults using a brain plasticity-based training program: a randomized, controlled study. *Proc. Natl. Acad. Sci. U.S.A.* 103, 12523–12528. doi: 10.1073/pnas.0605194103

Mårtensson, J., Eriksson, J., Bodammer, N. C., Lindgren, M., Johansson, M., Nyberg, L., et al. (2012). Growth of language-related brain areas after foreign language learning. *Neuroimage* 63, 240–244. doi: 10.1016/j.neuroimage.2012.06.043

McDowell, I., Xi, G., Lindsay, J., and Tierney, M. (2007). Mapping the connections between education and dementia. *J. Clin. Exp. Neuropsychol.* 29, 127–141. doi: 10.1080/13803390600582420

McMurtray, A., Saito, E., and Nakamoto, B. (2009). Language preference and development of dementia among bilingual individuals. *Hawaii Med. J.* 68, 223–226.

Mechelli, A., Crinion, J. T., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R. S., et al. (2004). Neurolinguistics: structural plasticity in the bilingual brain. *Nature* 431:757. doi: 10.1038/431757a

Mukadam, N., Sommerlad, A., and Livingston, G. (2017). The relationship of bilingualism compared to monolingualism to the risk of cognitive decline or dementia: a systematic review and meta-analysis. *J. Alzheimers Dis.* 58, 45–54. doi: 10.3233/JAD-170131

Ong, G., Sewell, D. K., Weekes, B., McKague, M., and Abutalebi, J. (2017). A diffusion model approach to analysing the bilingual advantage for the Flanker task: the role of attentional control processes. *J. Neurolinguistics* 43, 28–38. doi: 10.1016/j.jneuroling.2016.08.002

Ossher, L., Bialystok, E., Craik, F. I. M., Murphy, K. J., and Troyer, A. K. (2013). The effect of bilingualism on amnestic mild cognitive impairment. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 68, 8–12. doi: 10.1093/geronb/gbs038

Paap, K. R., Johnson, H. A., and Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex* 69, 265–278. doi: 10.1016/j.cortex.2015.04.014

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., and Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychol. Aging* 17, 299–320. doi: 10.1037/0882-7974.17.2.299

Park, D. C., Lodi-Smith, J., Drew, L. M., Haber, S., Hebrank, A., Bischof, G. N., et al. (2014). The impact of sustained engagement on cognitive function in older adults: the synapse project. *Psychol. Sci.* 25, 103–112. doi: 10.1177/0956797613499592

Peal, E., and Lambert, W. E. (1962). The relation of bilingualism to intelligence. *Psychol. Monogr. Gen. Appl.* 76, 1–23. doi: 10.1037/h0093840

Perani, D., and Abutalebi, J. (2015). Bilingualism, dementia, cognitive and neural reserve. *Curr. Opin. Neurol.* 28, 618–625. doi: 10.1097/WCO.0000000000000267

Perani, D., Farsad, M., Ballarini, T., Lubian, F., Malpetti, M., Fracchetti, A., et al. (2017). The impact of bilingualism on brain reserve and metabolic connectivity in Alzheimer's dementia. *Proc. Natl. Acad. Sci. U.S.A.* 114, 1690–1695. doi: 10.1073/pnas.1610909114

Perquin, M., Vaillant, M., Schuller, A. M., Pastore, J., Dartigues, J. F., Lair, M. L., et al. (2013). Lifelong exposure to multilingualism: new evidence to support cognitive reserve hypothesis. *PLOS ONE* 8:e62030. doi: 10.1371/journal.pone.0062030

Ramos, S., Fernández García, Y., Antón, E., Casaponsa, A., and Duñabeitia, J. A. (2017). Does learning a language in the elderly enhance switching ability? *J. Neurolinguistics* 43, 39–48. doi: 10.1016/j.jneuroling.2016.09.001

Raz, N. (2000). "Aging of the brain and its impact on cognitive performance," in *The Handbook of Aging and Cognition*, eds F. I. M. Craik and T. A. Salthouse (Mahwah, NJ: Lawrence Erlbaum Associates), 1–90.

Rodríguez-Fornells, A., Cunillera, T., Mestres-Missé, A., and de Diego Balaguer, R. (2009). Neurophysiological mechanisms involved in language learning in adults. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 3711–3735. doi: 10.1098/rstb.2009.0130

Salloway, S., Sperling, R., Fox, N. C., Blennow, K., Klunk, W., Raskind, M., et al. (2014). Two phase 3 trials of bapineuzumab in mild-to-moderate Alzheimer's disease. *N. Engl. J. Med.* 370, 322–333. doi: 10.1056/NEJMoa1304839

Sanders, A. E., Hall, C. B., Katz, M. J., and Lipton, R. B. (2012). Non-native language use and risk of incident dementia in the elderly. *J. Alzheimers Dis.* 29, 99–108. doi: 10.3233/JAD-2011-111631

Schroeder, S. R., and Marian, V. (2016). Cognitive consequences of trilingualism. *Int. J. Biling.* 21, 754–773. doi: 10.1177/1367006916637288

Schweizer, T. A., Ware, J., Fischer, C. E., Craik, F. I. M., and Bialystok, E. (2012). Bilingualism as a contributor to cognitive reserve: evidence from brain atrophy in Alzheimer's disease. *Cortex* 48, 991–996. doi: 10.1016/j.cortex.2011.04.009

Selkoe, D. J. (2012). Preventing Alzheimer's disease. *Science* 337, 1488–1492. doi: 10.1126/science.1228541

Smith, G. E., Housen, P., Yaffe, K., Ruff, R., Kennison, R. F., Mahncke, H. W., et al. (2009). A cognitive training program based on principles of brain plasticity: results from the improvement in memory with plasticity-based adaptive cognitive training (IMPACT) study. *J. Am. Geriatr. Soc.* 57, 594–603. doi: 10.1111/j.1532-5415.2008.02167.x

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

Staff, R. T., Murray, A. D., Deary, I. J., and Whalley, L. J. (2004). What provides cerebral reserve? *Brain* 127, 1191–1199. doi: 10.1093/brain/awh144

Stein, M., Federspiel, A., Koenig, T., Wirth, M., Strik, W., Wiest, R., et al. (2012). Structural plasticity in the language system related to increased second language proficiency. *Cortex* 48, 458–465. doi: 10.1016/j.cortex.2010.10.007

Stein, M., Winkler, C., Kaiser, A. C., and Dierks, T. (2014). Structural brain changes related to bilingualism: does immersion make a difference? *Front. Psychol.* 5:1116. doi: 10.3389/fpsyg.2014.01116

Stern, Y. (2012). Cognitive reserve in ageing and Alzheimer's disease. *Lancet Neurol.* 11, 1006–1012. doi: 10.1016/S1474-4422(12)70191-6

Stilwell, B. L., Dow, R. M., Lamers, C., and Woods, R. T. (2016). Language changes in bilingual individuals with Alzheimer's disease. *Int. J. Lang. Commun. Disord.* 51, 113–127. doi: 10.1111/1460-6984.12190

Tao, L., Marzecova, A., Taft, M., Asanowicz, D., and Wodniecka, Z. (2011). The efficiency of attentional networks in early and late bilinguals: the role of age of acquisition. *Front. Psychol.* 2:123. doi: 10.3389/fpsyg.2011.00123

Valenzuela, M. J., Jones, M., Wen, W., Rae, C., Graham, S., Shnier, R., et al. (2003). Memory training alters hippocampal neurochemistry in healthy elderly. *Neuroreport* 14, 1333–1337. doi: 10.1097/01.wnr.0000077548.91466.05

Valenzuela, M. J., and Sachdev, P. (2006). Brain reserve and dementia: a systematic review. *Psychol. Med.* 36, 441–454. doi: 10.1017/S0033291705006264

Valian, V. (2015a). Bilingualism and cognition. *Bilingualism* 18, 3–24. doi: 10.1017/S1366728914000522

Valian, V. (2015b). Bilingualism and cognition: a focus on mechanisms. *Bilingualism* 18, 47–50. doi: 10.1017/S1366728914000698

Veenstra, W. S., Huisman, M., and Miller, N. (2014). Age of acquisition and naming performance in Frisian-Dutch bilingual speakers with dementia. *Dement. Neuropsychol.* 8, 249–259. doi: 10.1590/S1980-57642014DN83000009

Verghese, J., Lipton, R. B., Katz, M. J., Hall, C. B., Derby, C. A., Kuslansky, G., et al. (2003). Leisure activities and the risk of dementia in the elderly. *N. Engl. J. Med.* 348, 2508–2516. doi: 10.1056/NEJMoa022252

Volkmar, F. R., and Greenough, W. T. (1972). Rearing complexity affects branching of dendrites in the visual cortex of the rat. *Science* 176, 1445–1447. doi: 10.1126/science.176.4042.1445

Ware, C., Damnee, S., Djabelkhir, L., Cristancho, V., Wu, Y., Benovici, J., et al. (2017). Maintaining cognitive functioning in healthy seniors with a technology-based foreign language program: a pilot feasibility study. *Front. Aging Neurosci.* 9:42. doi: 10.3389/fnagi.2017.00042

Weber, A., and Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *J. Mem. Lang.* 50, 1–25. doi: 10.1016/S0749-596X(03)00105-0

Whalley, L. J., Deary, I. J., Appleton, C. L., and Starr, J. M. (2004). Cognitive reserve and the neurobiology of cognitive aging. *Ageing Res. Rev.* 3, 369–382. doi: 10.1016/j.arr.2004.05.001

Wilson, R. S., Boyle, P. A., Yang, J., James, B. D., and Bennett, D. A. (2015). Early life instruction in foreign language and music and incidence of mild cognitive impairment. *Neuropsychology* 29, 292–302. doi: 10.1037/neu0000129

Woumans, E., Santens, P., Sieben, A., Versijpt, J., Stevens, M., and Duyck, W. (2015). Bilingualism delays clinical manifestation of Alzheimer's disease. *Bilingualism* 18, 568–574. doi: 10.1017/S136672891400087X

Woumans, E., Versijpt, J., Sieben, A., Santens, P., and Duyck, W. (2017). Bilingualism and cognitive decline: a story of pride and prejudice. *J. Alzheimers Dis.* 60, 1237–1239. doi: 10.3233/JAD-170759

Yang, J., Gates, K. M., Molenaar, P., and Li, P. (2015). Neural changes underlying successful second language word learning: an fMRI study. *J. Neurolinguistics* 33, 29–49. doi: 10.1016/j.jneuroling.2014.09.004

Yeung, C. M., St John, P. D., Menec, V., and Tyas, S. L. (2014). Is bilingualism associated with a lower risk of dementia in community-living older adults? Cross-sectional and prospective analyses. *Alzheimer Dis. Assoc. Disord.* 28, 326–332. doi: 10.1097/WAD.0000000000000019

Zahodne, L. B., Schofield, P. W., Farrell, M. T., Stern, Y., and Manly, J. J. (2014). Bilingualism does not alter cognitive decline or dementia risk among Spanish-speaking immigrants. *Neuropsychology* 28, 238–246. doi: 10.1037/neu0000014

Zou, L., Ding, G., Abutalebi, J., Shu, H., and Peng, D. (2012). Structural plasticity of the left caudate in bimodal bilinguals. *Cortex* 48, 1197–1206. doi: 10.1016/j.cortex.2011.05.022

Changes in White-Matter Connectivity in Late Second Language Learners: Evidence from Diffusion Tensor Imaging

Eleonora Rossi[1,2]*, Hu Cheng[3], Judith F. Kroll[2], Michele T. Diaz[4] and Sharlene D. Newman[3]

[1] Department of Psychology and Sociology, California State Polytechnic University, Pomona, Pomona, CA, United States,
[2] Department of Psychology, University of California, Riverside, Riverside, CA, United States, [3] Department of Psychology,
Indiana University, Bloomington, IN, United States, [4] Department of Psychology, Pennsylvania State University, University
Park, PA, United States

Morphological brain changes as a consequence of new learning have been widely established. Learning a second language (L2) is one such experience that can lead to rapid structural neural changes. However, still relatively little is known about how levels of proficiency in the L2 and the age at which the L2 is learned influence brain neuroplasticity. The goal of this study is to provide novel evidence for the effect of bilingualism on white matter structure in relatively proficient but late L2 learners who acquired the second language after early childhood. Overall, the results demonstrate a significant effect on white matter fractional anisotropy (FA) as a function of L2 learning. Higher FA values were found in a broad white matter network including the anterior thalamic radiation (ATR), the inferior fronto-occipital fasciculus (IFOF), the Uncinate Fasciculus (UF), and the inferior longitudinal fasciculus (ILF). Moreover, FA values were correlated with age of L2 acquisition, suggesting that learning an L2, even past childhood, induces neural changes. Finally, these results provide some initial evidence that variability in the age of L2 acquisition has important consequences for neural plasticity.

Keywords: diffusion tensor imaging, bilingualism, second language learning, neuroplasticity, age of acquisition

Learning a second language (L2) after a putative critical period for language learning (Long, 1990; Birdsong, 1999) is notably difficult, especially when the native language (L1) and the L2 are linguistically different. Past research on late L2 attainment suggesting mixed outcomes has been interpreted in different ways. One perspective proposes that late L2 representation and processing is hard-wired by maturational constraints and is fundamentally different than native language processing, especially when the grammatical structures of the two languages differ (e.g., Johnson and Newport, 1991; Weber-Fox and Neville, 1996; MacWhinney, 2005; Clahsen and Felser, 2006; Sabourin et al., 2006; Sabourin and Stowe, 2008). In contrast, processing-based accounts of L2 acquisition posit that native-like processing is possible for individuals who acquire an L2 after childhood, with some late learners acquiring a high level of L2 proficiency (e.g., McDonald, 2000; Birdsong and Molis, 2001; McLaughlin et al., 2010; Coughlin and Tremblay, 2012; Rossi et al., 2014). Other studies have shown that proficient late L2 speakers are also able to exploit cognitive resources that are central for on-line language processing (e.g., Hopp, 2010, 2014; Linck et al., 2014). Moreover, near native-like L2 processing has been correlated with immersion in the L2

environment, even when the experience was brief (Linck et al., 2009), suggesting that L2 processing is sensitive to variability in the frequency of usage and characteristics of L2 exposure (Ellis and Ogden, 2017).

The long-standing question of the nature of L2 representation and processing has also been extended to the realm of its neural underpinnings and has fueled a wealth of functional neuroimaging research with the goal of investigating if the functional neural networks underlying L2 processing are similar to the ones observed during native language processing, and to ask whether variables such as proficiency and age of acquisition (AoA) modulate the recruitment of those networks (see Li et al., 2014; García-Pentón et al., 2015; Luk and Pliatsikas, 2016 for recent reviews). Overall, functional evidence suggests that both languages are supported by similar cortical substrates even when the L2 is acquired relatively later in life, and that the recruitment of those networks is influenced by AoA (Perani et al., 1996, 1998; Wartenburger et al., 2003; Perani and Abutalebi, 2005) and also proficiency levels (Perani et al., 1998; Abutalebi et al., 2001). Very recently however, Xu et al. (2017) used multivariate pattern analysis (MVPA) to challenge the traditional single cortical mechanism hypothesis, proposing instead that the two languages might share the same neural substrate but may critically be supported by functionally independent networks. Critically, bilingualism and L2 learning lead also to the reorganization of neural areas that are not specifically related to language processing, but are involved in domain-general executive functions (Crinion et al., 2006; Li et al., 2014; Bialystok, 2017). The recruitment of domain-general brain areas such as the anterior cingulate cortex (Abutalebi et al., 2012), and subcortical structures such as the caudate (Abutalebi et al., 2008; Branzi et al., 2015) have been linked to mechanisms involved in language regulation, activation, and selection that are necessary in the face of ubiquitous co-activation of both languages, even when bilinguals intend to speak one language alone (e.g., Costa, 2005; Kroll et al., 2006). One prominent account proposes that for bilinguals to be able to successfully speak and control their languages, they engage a dynamic domain general neural network involving cortical and subcortical brain structures that allows them to resolve language competition to successfully select the intended language (Green, 1998; Abutalebi and Green, 2007; Green and Abutalebi, 2013).

Despite the wealth of research on the functional underpinnings of L2 processing, fewer studies have investigated the extent to which learning an L2 promotes *structural* brain changes. Early seminal research on neuroplasticity in animal models (Rosenzweig et al., 1962; Bennett et al., 1964; Diamond et al., 1964) demonstrated that the brain is not an immutable organ, but is pliable, and influenced by enriched environmental conditions and different task demands. Similarly, research on structural and morphological brain changes in the human brain have revealed that the brain is highly malleable and changes as a function of different types of skill learning. Neuroplastic changes in gray matter (GM) and white matter (WM) have been demonstrated across a vast array of skill and motor learning tasks (Draganski et al., 2004; Bengtsson et al., 2005), visual memory (Maguire et al., 2000), music practice (Skare et al.,

2005), and even higher-level meditation practices (Hernández et al., 2016).

Crucially, learning and juggling two languages constitute a prime example of new skill acquisition, especially when the L2 is learned past childhood and its acquisition is largely dependent on explicit learning mechanisms (Ullman, 2016). It is possible that *late* L2 learning in particular might be considered the perfect testbed to examine the effect of neuroplastic changes as a consequence of language learning. In fact, actively learning and mastering an L2, especially later in life might involve re-training and restructuring of a number of neural structures related to L2 language production, articulation, and language comprehension, potentially leading to greater neural changes especially during the most active learning phases (Xiang et al., 2015). Although, neuroplasticity may decrease across the life-span (Kennedy and Raz, 2009) resulting in smaller detectable changes after childhood, we hypothesize that adolescent or adult L2 language learning may be a sufficiently challenging task to elicit neural changes even in the face of reduced neuroplasticity. This idea resonates with the literature on desirable difficulties in learning, which proposes that L2 language learning and use is inherently taxing for the cognitive and neural system, but it is exactly that inherent difficulty that will produce long-term positive consequences for domain-general functions (Bjork and Kroll, 2015).

Evidence in favor of the neuroplastic effects of bilingualism is growing (Costa and Sebastián-Gallés, 2014). In a seminal study, Mechelli and colleagues demonstrated that bilinguals have greater GM density in the left inferior parietal lobule than monolingual controls (Mechelli et al., 2004; Della Rosa et al., 2013), and that the effect is modulated by AoA and proficiency, with earlier exposure to the L2 and higher L2 proficiency being positively correlated with higher GM. Similarly, greater GM density in the left inferior parietal gyrus (LIPG) has been reported in older bilingual adults (Abutalebi et al., 2015a), however with no correlations with AoA or proficiency. Differences between bilinguals and monolinguals in GM surface area and cortical thickness have also been shown in non-language related areas, with greater GM in the anterior cingulate cortex (Abutalebi et al., 2012; Felton et al., 2017). Finally, greater GM volume in bilinguals has been documented in several other areas, including the caudate nucleus (e.g., Grogan et al., 2012; Zou et al., 2012b), and putamen (Abutalebi et al., 2013a) which are subcortical areas that are important for language selection and control, both in non-pathological bilingual language processing (e.g., Abutalebi et al., 2007), and in the face of pathology (Green and Abutalebi, 2008). Increases in GM density in left IFG have also been found after a 5-month period of immersed L2 learning, suggesting again that L2 learning promotes fast neural restructuring (Stein et al., 2012).

Research on the neural changes promoted by bilingualism and L2 learning has also examined effects on white matter connectivity. To date however, even though the literature is rapidly growing, the majority of the research has examined simultaneous or early bilinguals who acquired their two languages during early childhood. For example, a study comparing early bilingual children to sequential bilingual

children (who learned the L2 at 3 years old) and monolingual children revealed that white matter microstructure (measured through fractional anisotropy, FA) in language-related bundles is positively modulated by bilingualism, and has provided evidence that the magnitude of the effect is dependent on AoA (Mohades et al., 2012, 2015). In these studies, Mohades and colleagues analyzed four WM tracts, including the left inferior frontal-occipital fasciculus (IFOF), the left arcuate fasciculus/superior longitudinal fasciculus (SLF), the WM bundle from the anterior part of the corpus callosum projecting to the orbital frontal cortex, and WM fibers from the anterior midbody of the corpus callosum to premotor and supplementary motor cortices. Their results showed that simultaneous bilinguals have higher FA in L-IFOF which is a ventral WM pathway that has been proposed to be central during spoken word recognition (Leclercq et al., 2010), and semantic processing (Duffau, 2008; Duffau et al., 2009; Martino et al., 2010). Mohades and colleagues also reported that sequential bilinguals had intermediate FA values between monolinguals and simultaneous bilinguals. They concluded that early bilingualism leads to neural adaptation in the human brain. In a follow up 2-year longitudinal study, Mohades et al. (2015) tracked simultaneous, and sequential bilingual children who were learning an L2. The results showed again higher FA values in IFOF for simultaneous bilinguals, but crucially sequential bilinguals showed an even greater change in IFOF over the course of the 2 years. The authors concluded that the degree of neural reshaping induced by bilingualism and L2 learning is partly dependent on AoA. Similar conclusions have been reported by Hämäläinen et al. (2017) who compared a group of early and sequential bilinguals. They analyzed mean FA, mean and radial diffusivity (MD and RD), and found that early bilingualism led to higher WM in the arcuate fasciculus, while sequential bilinguals showed greater WM connectivity in bilateral Inferior fronto-occipital fasciculus (IFOF), suggesting that different ages of L2 acquisition might determine what WM tracts might be shaped by language experience. Recent data has also revealed separate WM structural networks depending on different AoA but also proficiency levels, suggesting that brain changes might be differentially shaped by these two factors (Nichols and Joanisse, 2016). In sum, research on WM changes in early bilinguals has demonstrated that acquiring two languages from early childhood, or even learning an L2 relatively early during childhood has neuroplastic effects on both language specific and domain general WM pathways (Kousaie et al., 2017). Importantly, studies of WM changes in early bilingualism highlight that it is misleading to characterize AoA as a discrete variable, but rather that AoA should be understood as a continuum, as even within early acquisition, differences in AoA are correlated with different quantitative and qualitative effects.

To date, relatively few studies have investigated white-matter reorganization following L2 acquisition past early childhood. One such study investigated differences in WM structures between monolinguals and young adults who were late L2 learners (Pliatsikas et al., 2015). The L2 learners ($n = 20$) had a variety of languages as their L1s and had acquired English past the age of 10, but were classified as highly proficient English speakers. Participants were tested in the UK, thus immersed in their L2 environment. The TBSS results revealed higher FA values for the L2 group in the corpus callosum, including the genu, the body, and the anterior part of the splenium. Higher FA values were also found in left and right IFOF, bilateral uncinate fasciculi, and superior longitudinal fasciculi, all WM tracts that have been found to be modulated in early bilinguals. However, no correlational effects were found with length of immersion in the L2. The authors concluded that there is an effect of bilingualism on WM structures even when the L2 is learned past childhood. Importantly, the observed WM structures that have been identified for late bilinguals are similar to the ones that have been reported to be shaped by bilingualism in older adults (Luk et al., 2011), and also in early bilinguals (Mohades et al., 2012), suggesting that neural structures undergo neuroplastic changes as a consequence of L2 learning and bilingualism irrespective of the age at which the L2 is acquired. Similar neuroplastic changes in white matter have been reported in Spanish-English bilinguals who immigrated to the US in adulthood (Kuhl et al., 2016), and who learned English later in life (mean age = 19.4 years; range = 4.5–28.5 years). These speakers were immersed in their L2 environment at testing, and were recruited from the general population. The results reveled higher FA values in the bilateral anterior thalamic radiation (ATR), a bundle of fibers that are part of the internal capsule, and carry nerve fibers between the thalamus and the prefrontal cortex. Additionally, Kuhl and colleagues found a positive correlation between FA values and years of immersion in the L2, and with speaking abilities, suggesting that the degree of neural restructuring in ATR was proportional to L2 language experience.

However, other studies have reported contrary results to the ones presented above. For example, Cummine and Boliek (2013) tested adult Chinese–English bilinguals (mean age, 24.2; L2 AoA before the age of 5) and 11 English monolinguals (mean age, 28.5). The results showed significant decrease in FA for bilinguals as compared to monolinguals in the right inferior frontal-occipital fasciculus (IFOF), and in the superior portion of the right anterior thalamic radiation, and bilaterally in the inferior portion. These results are also in line with other studies that did report minimal differences between bilingual children and monolinguals (e.g., Mohades et al., 2012).

A similar approach to studying the effects of late L2 learning on WM has been taken in studies that have asked the question of what neural changes occur when learning happens during a relatively short but intensive program of language training. Mamiya et al. (2016) recruited 44 native college-age Chinese speakers who were enrolled in a 16-day upper level English course. They collected structural scans (DTI) between the 11th day of the course and 8 days after the course ended. For those participants who were tested before the end of the course, results showed a significant cluster of activation in the right and left SLF, and a positive correlation with the number of days in the course. The same study also revealed a marginally negative correlation between FA values in the right SLF, and days passed after the end of the immersion course. The authors concluded that there is a relationship between the diffusion properties of the brain and the length of immersion, suggesting that changes in white matter are rapid. Similar results were reported by

Schlegel et al. (2012) who tracked changes in white matter connectivity in a group of adult learners (mean age: 20.5) during a relatively longer 9-month intensive Chinese language course. Scans were acquired every month, and were compared to those of a comparable control group of individuals who did not attend any language course. Results showed a significant increase in FA values only for the learners in language-related WM tracts in the left hemisphere and in the genu of the corpus callosum, suggesting a strengthening of inter-hemispheric connections during L2 learning. Tract-based analysis also revealed that the learners group showed higher FA values in a number of tracts, some of which terminated in the left caudate nucleus which is implicated in language control (Green and Abutalebi, 2008), response selection (e.g., Branzi et al., 2015) and language-switching (Abutalebi et al., 2007). Similar results were found for a cohort of Japanese speakers who underwent 16 weeks of intensive English vocabulary training, while MRI scans were acquired before and after the training. Results revealed changes in right inferior frontal gyrus (IFG), arcuate fasciculus, and the pathways that connect IFG with the caudate nucleus (Hosoda et al., 2013). However, the observed WM changes reverted to baseline after 1 year, suggesting that neuroplastic changes might change depending on different demands. Similarly, Xiang et al. (2015) who tested a group of native German speakers who were enrolled in a 6-week intensive Dutch course while immersed in The Netherlands. Structural scans were acquired before and after the Dutch course. Results revealed a quick structural neural reorganization in connection with increasing L2 proficiency. A shift in hemispheric dominance was observed during early learning with greater FA values observed in the right arcuate fasciculus at early time points, which shifted back to the left with higher levels of L2 proficiency.

In sum, the recent literature suggests that WM pathways are modulated by L2 learning and bilingualism. However, evidence is still mixed regarding the relative contributions of proficiency and AoA, with data suggesting that that both proficiency and AoA play an important role in modulating those changes (e.g., Nichols and Joanisse, 2016; Hämäläinen et al., 2017). Regarding which WM pathways are most strongly impacted by bilingualism, a number of WM pathways have been highlighted as being frequently related to L2 learning and bilingualism. One such WM pathway is the SLF, a dorsal language network which connects posterior (superior temporal gyrus/Wernicke's area) and anterior (inferior frontal gyrus/ Broca's area) language cortices (Hickok and Poeppel, 2004, 2007). The IFOF instead, connects a ventral language network that includes Broca's area and posterior occipitotemporal regions, and also connects the anterior temporal lobe with the uncinate fasciculus (Anwander et al., 2007).

The goal of this study is to further examine the effects of L2 learning on WM in late L2 learners. To assess changes in WM we measured differences in fractional anisotropy (FA) in a group of monolingual speakers ($n = 24$) and a group of native English speaking late L2 learners of Spanish ($n = 24$) using tract-based spatial statistics (TBSS; Smith et al., 2006). FA can be used as an index of WM integrity, by reflecting the degree of anisotropy in water flow within the brain (Kunimatsu et al.,

2004). If late L2 learning promotes neural adaptation, we should observe differences in FA values between the L2 learners and the monolinguals in WM tracts that have been previously found to be positively affected by bilingualism. For example, the left inferior frontal-occipital fasciculus (IFOF; Mohades et al., 2012, 2015; Pliatsikas et al., 2015; see García-Pentón et al., 2015 for an extensive review) which is closely connected to the left ILF (Wakana et al., 2007), the uncinate fasciculus, which has been implicated in naming (Catani and Mesulam, 2008; Papagno, 2011), and found to be modulated by bilingualism (Hosoda et al., 2013; Qi et al., 2015). Moreover, if late L2 learning also affects domain-general brain networks, effects should be seen in cortical-subcortical WM fibers that have been proposed to be utilized during bilingual language selection and control, such as fibers that connect the IFG with the caudate (Tan et al., 2011; Hosoda et al., 2013).

An additional goal of this study was to contribute to the growing literature on how proficiency and AoA, as well as factors related to L2 use and experience, such as length of immersion in an L2 environment, contribute to the observed neural restructuring.

MATERIALS AND METHODS

Participants
Twenty-four monolingual English speakers (15 females), and 25 (20 females) native English speaking, late learners of Spanish participated in the study (age range: 18–27). All participants were recruited from the student population at Pennsylvania State University and all were right-handed. They were screened for safety, and contraindications to MRI scanning, in accordance with IRB requirements. None of them reported having been diagnosed with any neurological or reading disorder and all had normal or corrected-to normal visual acuity. All participants completed a language history questionnaire to assess their language history and skills. The results from the questionnaire showed that English monolingual speakers had no or minimal knowledge of a second language. L2 Spanish speakers were native speakers of English who learned Spanish as their second language later in life (average L2 acquisition age: 12 years). They all reported to be dominant speakers of English. Participants rated their L1 and L2 language knowledge using a scale from 1 to 10 (1 being the lowest and 10 being the highest score) for oral comprehension, oral production, reading and writing. They were paid for their participation and all study procedures were approved by the IRB at Penn State University.

Materials
As part of the testing battery, participants completed additional linguistic tasks that were designed to measure their proficiency in the L2 (Spanish). The language testing battery included a self-report language history questionnaire (reported in Appendix A) and a more objective grammar task. The primary task in the experiment was a picture naming task in English and Spanish that was part of an additional functional MRI study protocol that involved naming 6 runs of pictures (Rossi et al., in preparation). During this task, participants named a total of 144

items which were named in Spanish for the L2 learners group and in English for the English monolingual group. Stimuli consisted of images that were presented as line drawings, black and white photographs, or color photographs taken from 6 categories: animals, body parts, fruits, and vegetables, clothing, kitchen items, furniture. Within each category there were 16 items of each format for a total of 48 stimuli per category. Black and white and color photographs were identical except in color. The three formats were incorporated to allow for concept repetition, but minimize perceptually based priming. All images were 300×300 pixels and in bitmap image format. Across categories pictures were matched for frequency and imageability. All stimuli were presented using the Brain Logics MRI Digital Projection System, and experimental parameters were controlled via E-prime. Responses were recorded with an MR compatible microphone (Resonance Technologies, Northridge, CA). Examples of the stimuli are provided in **Figure 1**.

The grammar section of the Diploma de Español como Lengua Extranjera (DELE, Ministry of Education Culture Sport of Spain, 2006) was also administered to obtain an objective measure of grammatical knowledge in Spanish. Three sections of the DELE test were selected for this study. Participants completed the written text comprehension, the vocabulary and the grammar sections of the test. An example of the DELE test can be retrieved at: http://www.dele.org/. Finally, participants rated their L2 proficiency on a self-reported scale using a 0–10 scale, rating their language oral and written production and comprehension abilities. The full language history questionnaire is reported in Appendix A. The aggregate scores were calculated as follows: raw scores were standardized to z-scores and were summed together within each participant; then the resulting score was divided by the square root of the sum of the variances and covariances of all the subtests (Crocker and Algina, 1986; McMurray et al., 2010; Pivneva et al., 2012). These data are summarized in **Table 1**.

Imaging Pre-processing, Procedures, and Analysis

MRI scanning was completed on a Siemens 3.0 Tesla Magnetom Trio whole-body, human scanner (60 cm bore, 40 mT/m gradients, 200 T/m/s slew rate). An eight-channel head coil was used for Radio Frequency (RF) reception (Siemens Healthcare, Erlangen, Germany). Sagittal T-1 weighted localizer images were acquired and used to define a volume for high order shimming.
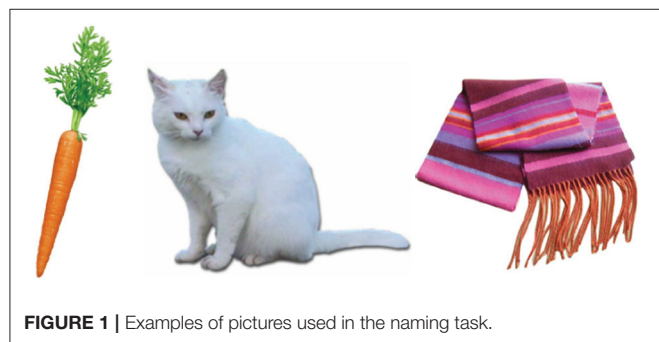

FIGURE 1 | Examples of pictures used in the naming task.

The anterior and posterior commissures were identified for slice selection and shimming. A semi-automated high-order shimming program was used to ensure global field homogeneity. High-resolution structural images were acquired using a 3D MP-RAGE pulse sequence (TR = 1,400 ms; TE = 2.01 ms; TI = 900 ms; FOV = 25.6 cm^2; flip angle = 9°; acceleration factor = 2; voxel size = $1 \times 1 \times 1$ mm; 160 contiguous slices).

Diffusion Tensor Imaging (DTI) data were collected using the following parameters: TR/TE = 6,500/93 ms, FOV = 240 mm, matrix = 128×128, 48 slices, slice thickness = 3 mm with 20% gap, averages = 2. iPAT factor = 2, phase partial Fourier = 6/8, 20 diffusion directions, b = 1,000 s/mm^2. DTI data were processed with FSL's FDT tool for eddy current correction and motion correction. Diffusion tensor was then computed using the tensor model to obtain FA values as inputs for TBSS analysis to examine the FA differences between monolinguals and English-Spanish bilinguals on the mean FA skeleton in FSL (Smith et al., 2004). The diffusion data were extracted first using BET (Smith, 2002). FA images were created by fitting a tensor model to the brain-extracted diffusion data using the FDT tool. FA's data were then are aligned into a common space using the non-linear registration tool FNIRT (Andersson et al., 2007a,b), which uses a b-spline representation of the registration warp field (Rueckert et al., 1999). Next, a mean FA image is created and thinned to create a mean FA skeleton, which represents the centers of all tracts common to the group. Each subject's aligned FA data is then projected onto this skeleton and permutation-based statistics of FA is conducted on all the voxels on the skeleton. In addition, regression analyses were performed to examine the relationship between FA and a number of behavioral and language usage measures, such as AoA, various measures of L2 proficiency (see below for details), and immersion in the L2.

RESULTS

The results showed a significant difference in FA between L2 learners and monolingual speakers in a broad network of WM tracts ($p < 0.05$, corrected). **Table 2** and **Figure 2** present the FA results from the group comparisons between the L2 group and the monolingual group. Sliced were selected each 5 mms and representative voxels were identified. For each WM cluster with significantly larger FA values for the L2 learners group, we report one representative voxel location in MNI152 standard space. Higher FA values were found for the L2 group in the anterior-posterior corona radiata, extending ventrally to the anterior and the retrolenticular portion of the internal capsule, up to the posterior thalamic radiation. More specifically, higher FA values for L2 learners were observed in the anterior and posterior corona radiata which represent a network of fibers that weaves through the internal capsule and that crosses with the fibers of the corpus callosum (CC), including WM tracts of the ATR, the inferior fronto-occipital fasciculus (IFOF), and the uncinate fasciculus (UF). Moreover, higher FA values within the ATR continued ventrally into the anterior limb. Greater FA values in L2 learners were found in the IFOF, ATR, and within bundles of the Inferior Longitudinal Fasciculus (ILF). Finally, greater

TABLE 1 | Demographic and L2 language measures for the L2 Spanish learners.

Gender	Average L2 self-rating	Picture naming accuracy Spanish	DELE score	L2 proficiency composite score	L2 Age of Acquisition (years)	Immersion time in L2 (months)
F	7.0	0.861	0.86	1.750	11	8
M	8.75	0.583	0.74	1.756	13	6
F	10.0	0.750	0.72	2.810	16	24
F	7.25	0.403	0.48	−0.978	14	NA
F	9.0	0.819	0.82	3.282	12	6
F	6.75	0.528	0.42	−1.206	9	NA
M	7.50	0.819	0.68	1.551	12	6
F	6.25	0.347	0.44	−2.092	12	3
M	5.25	0.514	0.54	−1.348	12	4
F	7.00	0.625	0.6	0.211	12	NA
M	7.45	0.625	0.58	0.421	2	0
M	6.25	0.569	0.62	−0.333	12	0
F	6.00	0.625	0.42	−1.240	5	20
F	7.00	0.708	0.44	−0.363	13.5	0
M	7.13	0.458	0.52	−0.688	12	4
F	6.58	0.181	0.36	−2.642	8	0
F	6.25	0.278	0.28	−3.002	12	0
M	6.75	0.319	0.46	−1.512	13.5	0
F	8.5	0.806	0.56	1.529	19	22
M	7.5	0.986	0.64	1.685	14	5
M	8.5	0.722	0.58	1.408	13	1
F	6.75	0.283	0.3	−2.467	16	2
F	8.00	0.222	0.38	−1.231	13	0
F	8.25	0.250	0.45	−0.755	12	8.5
F	5.00	0.501	0.48	−1.785	14	20
Avg.	7.22	0.55	0.53	−0.20	12.08	6.34
SD	1.17	0.22	0.15	1.75	3.46	7.83

NA: Data not provided in questionnaire.

FA values for L2 learners were found in the posterior thalamic radiation which has connections to ILF and IFOF WM tracts. Monolinguals did not show significantly higher FA values than bilinguals in any region.

In order to investigate whether FA values were correlated with measures of L2 language acquisition, proficiency, and length of immersion in the L2 environment, the mean FA from the voxels showing a significant difference between monolinguals and bilinguals was correlated with measures of, *L2 AoA*, *L2 proficiency* measured independently through the DELE grammar score, self-proficiency reports, a naming task, and a proficiency composite score (see Methods section for details), and *L2 immersion* measured in months. Results showed a significant correlation between FA and AoA ($r = -0.46$; $p = 0.02$), and between FA values and a normalized index of AoA (AoA/years of speaking the L2; $r = -0.465$; $p = 0.02$) which was calculated to normalize AOA values relative to the number of years participants had been speaking Spanish due to variation in the age of the participants (**Figure 3**). There were no significant correlations or trends between FA and proficiency, or FA and length of immersion in the L2 ($r = -0.21$; $p = 0.31$). Note that given that not all participants mentioned when they returned to the US from their study abroad experience, we do not have a precise metric to calculate time in the L1 environment after immersion in the L2. However, from the available data the minimum time elapsed from returning to testing was 4 months. For all the remaining participants (who provided that information) they all returned to the US more than 1 year before testing. Additionally, the correlation analysis was run excluding participants who did not report the length of their stay abroad in the language history questionnaire. We reasoned that part of why we fail to find a significant effect is that the distribution of the amount of time spent abroad was not sufficient to show an effect. Even though there was variability in the number of months spent abroad, there was a significant portion of participants who did not study abroad.
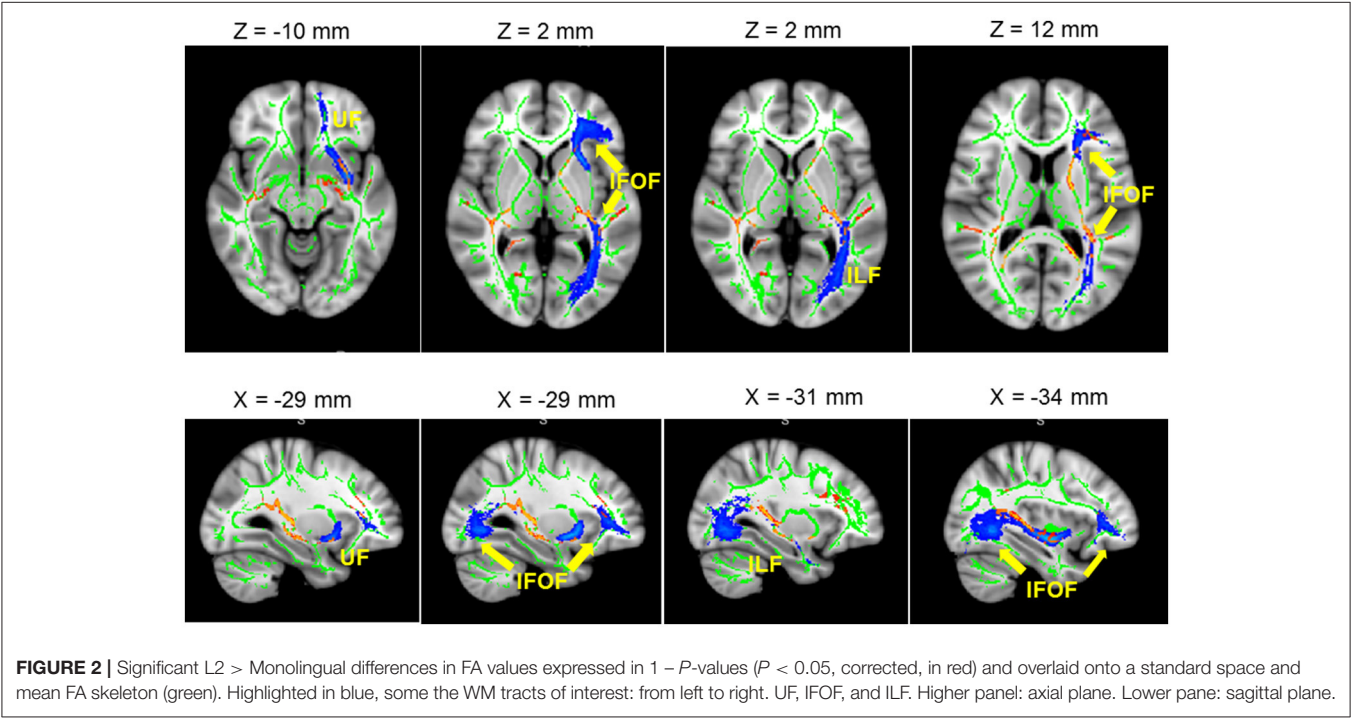
DISCUSSION AND CONCLUSIONS

The goal of this study was to investigate structural changes in WM related to L2 acquisition, especially when the L2 is acquired relatively later in life after a putative sensitive period for L2 learning (Long, 1990; Birdsong, 1999). We also asked whether observed changes were modulated by factors such as

TABLE 2 | WM clusters with significantly larger FA values for the L2 learners group.

x, y, z (Representative voxel)	Hem.	Brain area	WM tracts	p-value
−22, +18, +17	Left	Anterior Corona Radiata	ATR IFOF UF	p = 0.042
−27, −43, 22	Left	Posterior Corona Radiata	ATR IFOF	p = 0.032
−19, +4, +12 −19, +17, +2	Left	Internal Capsule: anterior limb	ATR	p = 0.042 p = 0.042
−32, −30, +7	Left	Internal Capsule: retrolenticular portion	IFOF	p = 0.034
−33, −27, +2	Left	Internal Capsule: retrolenticular portion	IFOF ILF ATR	p = 0.034
−33, −39, +12	Left	Posterior Thalamic Radiation	IFOF ILF	p = 0.032
38, −32, +2	Right	Internal Capsule: retrolenticular portion	IFOF ILF	p = 0.042
36, −33, +7	Right	Retrolenticular part of the internal capsule	IFOF ILF	p = 0.040
33, −34, +12	Right	Internal Capsule: retrolenticular portion	IFOF	p = 0.042

For each cluster, we report one representative voxel location in MNI space. ATR, Anterior thalamic radiation; IFOF, Inferior fronto-occipital fasciculus; UF, Uncinate Fasciculus; ILF, Inferior longitudinal fasciculus.

FIGURE 2 | Significant L2 > Monolingual differences in FA values expressed in 1 − P-values (P < 0.05, corrected, in red) and overlaid onto a standard space and mean FA skeleton (green). Highlighted in blue, some the WM tracts of interest: from left to right. UF, IFOF, and ILF. Higher panel: axial plane. Lower pane: sagittal plane.

AoA, proficiency, and language-use measures such as length of immersion in the L2 environment. WM fractional anisotropy was analyzed using TBSS and results were compared between a group of English-speaking, late L2 learners of Spanish and a group of monolingual English speakers.

The results revealed differences in WM FA between the two groups. L2 learners showed higher FA values in a number of WM tracts in the left hemisphere, including WM tracts of the

ATR, the IFOF, the uncinate fasciculus (UF), and the ILF. These results are in line with a number of studies that have reported adaptive WM changes in similar tracts in early (e.g., Mohades et al., 2012, 2015) and late bilingualism (e.g., Pliatsikas et al., 2015, 2017). The data we report supports the growing body of literature proposing that bilingualism and L2 learning promote not only functional but also structural neural adaptation (Li et al., 2014; Bialystok, 2017). Similar to the research conducted by Pliatsikas

FIGURE 3 | Correlation between mean FA and L2 normalized AoA $r = -0.465$; $p = 0.02$.
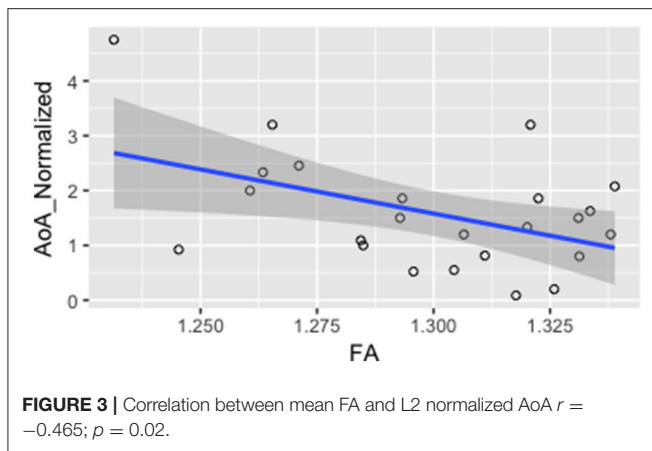
and colleagues (Pliatsikas et al., 2015, 2017), our study examined the effects of L2 learning on WM structure in adult speakers who learned the L2 later in life (average AoA = 12.1) and were therefore not early bilinguals. However, unlike the speakers in Pliatsikas et al. (2015, 2017), the L2 learners in our study were not immersed in their L2 environment at testing, but were immersed in their native language environment (English). This factor may account for some of the differences observed between our results and Pliatsikas et al.'s results. We will further discuss potential reasons for some of the observed differences between our results and Pliatsikas et al.'s results.

The results revealed higher FA values for L2 learners in the anterior and posterior corona radiata. This WM tract has been previously found to be part of a network implicated during simultaneous interpretation (Hervais-Adelman et al., 2014), and lesions in this region have been shown to lead to word retrieval problems in productive aphasia (Schnur et al., 2006) suggesting its importance for lexical retrieval. The present data are therefore in line with previous results on the recruitment of this WM tract during high-performance bilingual language processing (Hervais-Adelman et al., 2014). Our findings moreover reveal that this WM fiber tract is also implicated in lower performing bilinguals, suggesting that L2 learning also leads to neuro-adaptive changes in WM tracts that are at play during high-performing bilinguals, such as in simultaneous interpreters. The corona radiata is also one of the WM regions that have shown a consistent decline in FA in non-pathological aging (Kennedy and Raz, 2009). As such, these data also suggest that even late bilingualism may contribute increase neural connectivity in non-pathological aging (Kennedy and Raz, 2009), as well as in pathological brain decline (e.g., Luk et al., 2011; Abutalebi et al., 2015b; Perani and Abutalebi, 2015). Late L2 learning may strengthen neural pathways, including those that are most sensitive to age-related neural decline.

Across all of the identified WM regions, greater FA values in L2 learners were observed within the IFOF extending ventrally into the internal capsule, and the posterior thalamic radiation. According to recent language neural models, the IFOF represents a large ventral pathway implicated in language processing that connects the inferior frontal gyrus (BA45/47) with the superior temporal gyrus (BA22), the inferior parietal cortex (BA39), and the occipital cortex; all have been implicated in language comprehension and have been proposed to be central to semantics and spoken word recognition (Leclercq et al., 2010), and semantic processing in general (Duffau, 2008; Martino et al., 2010; see López-Barroso et al., 2013 for null results). Crucially, similar FA increase in the IFOF has been reported in a number of bilingual studies including in children (Mohades et al., 2012, 2015), younger adults (Pliatsikas et al., 2015), and in older bilingual adults (e.g., Luk et al., 2011; Gold et al., 2013 for null results FA values in IFOF in older adults), suggesting that bilingualism might contribute to boosting neural reserve, that might be accumulated throughout the life-span. The present results contribute to the evidence that bilingualism modulates the WM pathways implicated in language processing.

These data also showed higher FA values for L2 learners in the ATR, a bundle of fibers that weaves through the left anterior corona radiata and the anterior and retrolenticular limb of the internal capsule. ATR is a WM tract implicated in lexico-semantic processing by connecting a distributed language network in temporal, parietal, and frontal cortices (e.g., Han et al., 2013; Mirman et al., 2015). Disruption of ATR gives rise to semantic progressive primary aphasia (Han et al., 2013), and auditory verbal hallucinations (AVH) in schizophrenia are negatively correlated with FA values in the ATR (Curčić-Blake et al., 2015), highlighting its role during speech processing (Curčić-Blake et al., 2012). Although the literature has reported some null effects for this WM tract for bilinguals (e.g., Cummine and Boliek, 2013) significant differences within the ATR have been reported in older bilingual adults (Luk et al., 2011) and in children (Mohades et al., 2012). Higher FA values in bilateral ATR have also been found in Spanish-English bilinguals who emigrated to the US in adulthood (Kuhl et al., 2016), and who learned English later in life (mean age = 19.4 years; range = 4.5–28.5 years), and who were immersed in their L2 environment at testing. Additionally, Kuhl and colleagues also found a positive correlation between FA values in the ATR and years of immersion in the L2, and speaking abilities, suggesting that the degree of neural restructuring in ATR was proportional to L2 language experience. Overall, the present data corroborate previous results in showing higher FA values in bilinguals (early and late) for WM tracts that connect a distributed language network.

Finally, in line with our predictions, the results demonstrate higher FA values for L2 learners in the Uncinate Fasciculus, which connects inferior frontal and anterior temporal regions, and more ventrally in the ILF which connects anterior and posterior temporal regions. Although, the past literature is not conclusive regarding WM changes in the UF in bilingualism (García-Pentón et al., 2014; Grundy et al., 2017 for a recent review), the results we have reported are in line with a number of previous studies that have reported higher FA values in the UF in young late L2 learners (Pliatsikas et al., 2015), and in older bilingual adults (Luk et al., 2011). Anterior WM pathways such as the UF are involved in language production (Roelofs, 2014), aspects of syntactic processing (Friederici et al., 2006; Duffau et al., 2009; but see Teichmann et al., 2015 for a counter proposal), new word learning and consolidation (López-Barroso et al., 2013; Ripollés

et al., 2014), semantic performance in healthy older adults (de Zubicaray et al., 2011), and primary progressive aphasia (Han et al., 2013). The current results suggest that bilingualism and L2 learning enhance the utilization of those pathways to process both the L1 and the L2, possibly leading to higher neural integrity. Similarly, the ILF connects a ventral language network within temporal and posterior occipitotemporal regions to inferior frontal regions via the UF (Anwander et al., 2007). Consistent with our results, a number of WM connectivity studies have identified changes in the ILF in bilinguals (e.g., Hosoda et al., 2013), which have also been positively correlated with AoA (Nichols and Joanisse, 2016), supporting the idea that ILF pathways are strengthened through L2 processing.

Taken together, the data we have reported support the predictions based on previous WM studies in bilingualism, strengthening the general view that learning, and mastering a second language, even later in life, results in neuroplastic changes. Unlike past studies, the present study did not show a main effect in the corpus callosum (CC). Contrary to what we observed, a number of studies on WM across the life-span have reported effects of bilingualism on the CC (e.g., Luk et al., 2011; Schlegel et al., 2012; García-Pentón et al., 2015; Pliatsikas et al., 2015; but see also Cummine and Boliek, 2013, a study which reports no effects in the CC). One possible factor that might contribute to the absence of a clear effect in CC is our population, which included participants who were immersed in the L1, English, and not in the L2. It is also plausible to think that the absence of changes in CC in our data could be related to the acquisition of the L2 through more formal L2 instruction, rather than through pure immersion. Even though it is plausible to think that the bilingual speakers tested in Pliatsikas et al.'s (2015) did learn their L2 also through some formal learning either before moving to the UK, or while in the UK, they were immersed in the L2 environment at testing. Similarly, most of the studies that report effects in the CC tested participants who were either immersed in their L2 (e.g., Luk et al., 2011; Pliatsikas et al., 2015), or who were engaged in an intensive L2 language training (Schlegel et al., 2012). Again, because our sample of L2 speakers was immersed in their L1 environment, the reported results might, if anything, provide an underestimation of the effects that might be observed under conditions of immersion in the L2 (as for the sample of bilinguals in Pliatsikas et al., 2015). Moreover, our sample was more homogenous relative to Pliatsikas et al.'s relative to the L1. The speakers in the present study were all native speakers of English while Pliatsikas et al.'s bilinguals L1 language background varied. It could be hypothesized that variability in the L1, and linguistic distance between the L1 and the L2 may play a role in engaging neural pathways differentially. Future studies will need to address the question of how immersion in different linguistic environments, and variability in linguistic properties between the L1 and the L2 might modulate these effects.

An additional goal of this research was to understand how WM changes might be correlated with different measures of L2 learning and experience, including AoA, L2 proficiency levels, and length of immersion in the L2. First, the results showed a correlation between FA values and AoA, in line with previous studies on simultaneous and sequential bilingual children

(Mohades et al., 2012, 2015). The present data demonstrate that structural changes occur as a consequence of L2 learning, even when the L2 is acquired past childhood. Our data therefore demonstrate that L2 learning promotes neural restructuring. What is novel about these results is the observation that AoA might still be an important factor to explain variation in the observed structural changes, even when the L2 is learned past childhood. However, rather than reasoning in terms of AoA, which is a non-dynamic measure, it is tempting to propose that AoA should be rather interpreted more as a dynamic measure that could encompass length of time spent engaging in L2 learning.

The results did not reveal any correlation between mean FA values and proficiency in the L2. Previous results on the effects of proficiency on structural changes in L2 learners are at best mixed, and do not yet provide a clear picture of how variability in L2 proficiency affects structural changes. While some research has highlighted how structural changes can be correlated with increasing proficiency, and rate of successful L2 learning (e.g., Schlegel et al., 2012; Stein et al., 2012; Hosoda et al., 2013), other studies have failed to find a significant correlation between WM changes and proficiency (e.g., Stein et al., 2014). A number of factors could play a role in the lack of a correlation between structural csdhanges and proficiency across studies. First, proficiency is often measured in different ways across studies, and most importantly, those measures rely at times on subjective measures of proficiency only, and not on objective measures of performance. Additionally, AoA and proficiency are often highly conflated, preventing clear identification of their relative contributions. Finally, some of the studies that report a significant positive effect of proficiency on WM plasticity are training paradigms. For example, Hosoda et al. (2013) exposed participants to an intensive vocabulary learning regime, and they report that after intensive training changes in WM pathways such as the IFGop-caudate and IFGop-STG/SMG pathways were positively correlated with learning success. In this case however, it could be argued that changes in proficiency levels are an outcome by itself, rather than a predictor.

The participants tested in our study were not necessarily exposed to active L2 training. They were intermediate learners of Spanish who were recruited based on L2 Spanish knowledge, but who might or might have not been actively involved in Spanish learning at the moment of testing. We reason that the fact that there was variability in terms active exposure to L2 at testing might have contributed to the absence of correlation between FA and proficiency. Moreover, the variability in their L2 proficiency was pretty limited (mean accuracy on the picture naming in Spanish = 0.55; SD = 0.22), and this factor might also have undermined the likelihood to find a significant correlation between FA and L2 proficiency. Overall, future research should address variability in how proficiency is measured. More similar measures of L2 proficiency across studies would allow to create a common basis for measuring and comparing L2 proficiency across studies.

Evidence on the role of immersion in a naturalistic L2 environment as a catalyst for neural change is still mixed (Stein

et al., 2014; Pliatsikas and Chondrogianni, 2015). Changes in GM have been reported in simultaneous bilinguals (Burgaleta et al., 2016), in relatively early sequential trilinguals (Abutalebi et al., 2013b), and in later bilinguals (Zou et al., 2012a; DeLuca and Pliatsikas, 2017; Pliatsikas et al., 2017), and WM changes have been observed in immersed L2 learners in left inferior frontal-occipital fasciculus (L-IFOF) and SLF (Stein et al., 2014). The few studies that tested L2 learners who immersed themselves in the L2 environment past childhood mostly fail to observe a correlation between FA and length of immersion (Pliatsikas et al., 2015; DeLuca and Pliatsikas, 2017). Interestingly, a recent reanalysis of Pliatsikas et al.'s (2015) data using diffusion MRI connectometry and correlation analysis (Rahmani et al., 2017) revealed increased connectivity in corpus callosum (CC), arcuate fasciculus (AF), and left IFOF of sequential bilingual adults, and reported positive association with language immersion period, and showed that FA of all of the significant fibers from connectometry analysis, had direct correlation with the duration of immersion period of bilinguals. Our TBSS data does not show any correlation between FA and length of immersion in the L2 environment. One possible explanation for the absence of an effect is the small variability in length of immersion in the L2 (mean length in months: 6.64; range = 0–24), and the fact that all the L2 learners were immersed in their L1 environment at testing (contra Cummine and Boliek, 2013; Kuhl et al., 2016; DeLuca and Pliatsikas, 2017). Cummine and Boliek's (2013) study is likely to be the one with the most similar participants characteristics to ours. They tested adult Chinese–English bilinguals (mean age, 24.2; L2 AoA before the age of 5) and 11 English monolinguals (mean age, 28.5). Crucially however, in Cummine and Boliek, bilingual participants were immersed in the L2 (English), and no specific analyses were reported to correlate those results with length of immersion in the L2 environment. Our participants instead were immersed in the L1 environment at time of testing, and time since returning from being immersed in an L2 environment varied extensively. These factors might account for differences between the results reported by Cummine and Boliek and our results. In sum, given the relatively scarce literature on the potential neuroplastic effects of L2 immersion in late L2 bilinguals, a clear conclusion about the role of immersion will await future research.

To conclude, the present study reveals that L2 learning has the potential to shape the WM networks underlying language processing, even when the L2 is learned after childhood. Given the growing literature suggesting that L2 learning, and long-life experience with the L2 can lead to cognitive, and neural changes which might confer cognitive protection in healthy and pathological aging (inter alia Luk et al., 2011; Alladi et al., 2013; Gold et al., 2013; Bak et al., 2014a,b; Grady et al., 2015; Olsen et al., 2015), it is tempting to proposes that learning a second language throughout the life-span, even during adulthood should become one experiential form of continuous learning available to everyone.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the PSU IRB board with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the PSU IRB committee

AUTHOR CONTRIBUTIONS

ER: Designed the task, collected the data, contributed to the data analysis, and wrote the manuscript; HC: Performed the data analysis, and contributed to the preparation of the manuscript; JK: Designed the task, and contributed to the preparation of the manuscript; MD: Contributed to design the task, contributed to collected the data, and contributed to the preparation of the manuscript; SN: Contributed to collected the data, contributed to the data analysis, and contributed to the preparation of the manuscript.

FUNDING

REFERENCES

Abutalebi, J., Annoni, J. M., Zimine, I., Pegna, A. J., Seghier, M. L., Lee-Jahnke, H., et al. (2008). Language control and lexical competition in bilinguals: an event-related fMRI study. *Cereb. Cortex* 18, 1496–1505. doi: 10.1093/cercor/bhm182

Abutalebi, J., Brambati, S. M., Annoni, J.-M., Moro, A., Cappa, S. F., and Perani, D. (2007). The neural cost of the auditory perception of language switches: an event-related functional magnetic resonance imaging study in bilinguals. *J. Neurosci.* 27, 13762–13769. doi: 10.1523/JNEUROSCI.3294-07.2007

Abutalebi, J., Canini, M., Della Rosa, P. A., Green, D. W., and Weekes, B. S. (2015a). The neuroprotective effects of bilingualism upon the inferior parietal lobule: a structural neuroimaging study in aging Chinese bilinguals. *J. Neurolinguist.* 33, 3–13. doi: 10.1016/j.jneuroling.2014.09.008

Abutalebi, J., Cappa, S. F., and Perani, D. (2001). The bilingual brain as revealed by functional neuroimaging. *Bilingual. Lang. Cogn.* 4, 179–190. doi: 10.1017/S136672890100027X

Abutalebi, J., Della Rosa, P. A., Ding, G., Weekes, B., Costa, A., and Green, D. W. (2013a). Language proficiency modulates the engagement of cognitive control areas in multilinguals. *Cortex* 49, 905–911. doi: 10.1016/j.cortex.2012.08.018

Abutalebi, J., Della Rosa, P. A., Green, D. W., Hernandez, M., Scifo, P., Keim, R., et al. (2012). Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cereb. Cortex* 22, 2076–2086. doi: 10.1093/cercor/bhr287

Abutalebi, J., and Green, D. (2007). Bilingual language production: the neurocognition of language representation and control. *J. Neurolinguist.* 20, 242–275. doi: 10.1016/j.jneuroling.2006.10.003

Abutalebi, J., Guidi, L., Borsa, V., Canini, M., Della Rosa, P. A., Parris, B. A., et al. (2015b). Bilingualism provides a neural reserve for aging populations. *Neuropsychologia* 69, 201–210. doi: 10.1016/j.neuropsychologia.2015.01.040

Abutalebi, J., Rosa, P. A. D., Castro Gonzaga, A. K., Keim, R., Costa, A., Perani, D. (2013b). The role of the left putamen in multilingual language production. *Brain Lang.* 125, 307–315. doi: 10.1016/j.bandl.2012.03.009

Alladi, S., Bak, T. H., Duggirala, V., Surampudi, B., Shailaja, M., and Shukla, A. K., et al. (2013). Bilingualism delays age at onset of dementia,

independent of education and immigration status. *Neurology* 81, 1938–1944. doi: 10.1212/01.wnl.0000436620.33155.a4

Andersson, J. L., Jenkinson, M., and Smith, S. (2007a). *Non-Linear Registration, aka Spatial Normalisation FMRIB Technical Report TR07JA2*. FMRIB Analysis Group of the University of Oxford.

Andersson, J. L., Jenkinson, M., and Smith, S. (2007b). *Non-Linear Optimisation. FMRIB Technical Report TR07JA1*. University of Oxford FMRIB Centre, Oxford.

Anwander, A., Tittgemeyer, M., von Cramon, D. Y., Friederici, A. D., and Knösche, T. R. (2007). Connectivity-based parcellation of Broca's area. *Cereb. Cortex* 17, 816–825. doi: 10.1093/cercor/bhk034

Bak, T. H., Nissan, J. J., Allerhand, M. M., and Deary, I. J. (2014a). Does bilingualism influence cognitive aging? *Ann. Neurol.* 75, 959–963. doi: 10.1002/ana.24158

Bak, T. H., Vega-Mendoza, M., and Sorace, A. (2014b). Never too late? An advantage on tests of auditory attention extends to late bilinguals. *Front. Psychol.* 5:485. doi: 10.3389/fpsyg.2014.00485

Bengtsson, S. L., Zoltan, N., Skare, S., Forsman, L., Hans, F., and Fredrik, U. (2005). Extensive piano practicing has regionally specific effects on white matter development. *Nat. Neurosci.* 8, 1148–1151. doi: 10.1038/nn1516

Bennett, E. L., Diamond, M. C., Krech, D., and Rosenzweig, M. R. (1964). Chemical and anatomical plasticity of brain. *Science* 146, 610–619.

Bialystok, E. (2017). The bilingual adaptation : how minds accommodate experience. *Psychol. Bull.* 143, 233–262. doi: 10.1037/bul0000099

Birdsong, D. (1999). *Second Language Acquisition and The Critical Period Hypothesis*. Mahwah, NJ: Routledge.

Birdsong, D., and Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *J. Mem. Lang.* 44, 235–249. doi: 10.1006/jmla.2000.2750

Bjork, R. A., and Kroll, J. F. (2015). Desirable difficulties in vocabulary learning. *Am. J. Psychol.* 128, 241–252. doi: 10.5406/amerjpsyc.128.2.0241

Branzi, F. M., Della Rosa, P. A., Canini, M., Costa, A., and Abutalebi, J. (2015). Language control in bilinguals: monitoring and response selection. *Cereb. Cortex* 26, 2367–2380. doi: 10.1093/cercor/bhv052

Burgaleta, M., Sanjuán, A., Ventura-Campos, N., et al. (2016). Bilingual- ism at the core of the brain. Structural differences between bilinguals and monolinguals revealed by subcortical shape analysis. *Neuroimage* 125, 437–445. doi: 10.1016/j.neuroimage.2015.09.073

Catani, M., and Mesulam, M. (2008). The arcuate fasciculus and the disconnection theme in language and aphasia: history and current state. *Cortex* 44, 953-961. doi: 10.1016/j.cortex.2008.04.002

Clahsen, H., and Felser, C. (2006). Grammatical processing in language learners. *Appl. Psycholinguist.* 27, 3–42. doi: 10.1017/S0142716406060024

Costa, A. (2005). "Lexical access in bilingual production," in *Handbook of Bilingualism: Psycholinguistic Approaches*, eds J. F. Kroll and A. M. B. de Groot (New York, NY: Oxford University Press), 308–325.

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

Coughlin, C. E., and Tremblay, A. (2012). Proficiency and working memory based explanations for nonnative speakers' sensitivity to agreement in sentence processing. *Appl. Psycholinguist.* 34, 615–6143. doi: 10.1017/S0142716411000890

Crinion, J., Turner, R., Grogan, A, Hanakawa, T., Noppeney, U., Devlin, J. T., et al. (2006). Language control in the bilingual brain. *Science* 312, 1537–1540. doi: 10.1126/science.1127761

Crocker, L., and Algina, J. (1986). *Introduction to Classical and Modern Test Theory*. Fort Worth, TX: Holt, Rinehart and Winston.

Cummine, J., and Boliek, C. A. (2013). Understanding white matter integrity stability for bilinguals on language status and reading performance. *Brain Struct. Funct.* 218, 595–601. doi: 10.1007/s00429-012-0466-6

Curčić-Blake, B., Liemburg, E., Vercammen, A., Swart, M., Knegtering, H., Bruggeman, R., and Aleman, A. (2012). When Broca goes uninformed: reduced information flow to Broca's area in schizophrenia patients with auditory hallucinations. *Schizophr. Bull.* 39, 1087–1095. doi: 10.1093/schbul/sbs107

Curčić-Blake, B., Nanetti, L., van der Meer, L., Cerliani, L., Renken, R., Pijnenborg, G. M., and Aleman, A. (2015). Not on speaking terms: hallucinations and structural network disconnectivity in schizophrenia. *Brain Struct. Funct.* 220, 407–418. doi: 10.1007/s00429-013-0663-y

de Zubicaray, G. I., Rose, S. E., and McMahon, K. L. (2011). The structure and connectivity of semantic memory in the healthy older adult brain. *Neuroimage* 54, 1488–1494. doi: 10.1016/j.neuroimage.2010.08.058

Della Rosa, P. A., Videsott, G., Borsa, V. M., Canini, M., Weekes, B. S., Franceschini, R., et al. (2013). A neural interactive location for multilingual talent. *Cortex* 49, 605–608. doi: 10.1016/j.cortex.2012.12.001

DeLuca, V. and Pliatsikas C. (2017). "Structural effects of linguistic immersion on the bilingual brain: a longitudinal study," in *Poster Presented at the 2017 International Symposium on Bilingualism*, Limerick.

Diamond, M. C., Krech, D., and Rosenzweig, M. R. (1964). The effects of an enriched environment on the histology of the rat cerebral cortex. *J. Comp. Neurol.* 123, 111–119. doi: 10.1002/cne.901230110

Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U. andMay, A. (2004). Neuroplasticity: changes in grey matter induced by training. *Nature* 427, 311–312. doi: 10.1038/427311a

Duffau, H. (2008). The anatomo-functional connectivity of language revisited. New insights provided by electrostimulation and tractography. *Neuropsychologia* 46, 927–934. doi: 10.1016/j.neuropsychologia.2007.10.025

Duffau, H., Gatignol, P., Moritz-Gasser, S., and Mandonnet, E. (2009). Is the left uncinate fasciculus essential for language? *J. Neurol.* 256, 382–389. doi: 10.1007/s00415-009-0053-9

Ellis, N. C., and Ogden, D. C. (2017). Thinking about multiword constructions: usage-based approaches to acquisition and processing. *Top. Cogn. Sci.* 9, 604–620. doi: 10.1111/tops.12256

Felton, A., Vazquez, D., Ramos-Nuñez, A. I., Greene, M. R., MacBeth, A., Hernandez, A. E., et al. (2017). Bilingualism in fluences structural indices of interhemispheric organization. *J. Neurolinguist.* 42, 1–11. doi: 10.1016/j.jneuroling.2016.10.004

Friederici, A. D., Bahlmann, J., Heim, S., Schubotz, R. I., and Anwander, A. (2006). The brain differentiates human and non-human grammars: functional localization and structural connectivity. *Proc. Natl. Acad. Sci. U.S.A.* 103, 2458–2463. doi: 10.1073/pnas.0509389103

García-Pentón, L., Fernández García, Y., Costello, B., Duñabeitia, J. A., and Carreiras, M. (2015). The neuroanatomy of bilingualism: how to turn a hazy view into the full picture. *Lang. Cogn. Neurosci.* 3798, 1–25. doi: 10.1080/23273798.2015.1068944

García-Pentón, L., Pérez Fernández, A., Iturria-Medina, Y., Gillon-Dowens, M., and Carreiras, M. (2014). Anatomical connectivity changes in the bilingual brain. *Neuroimage* 84, 495–504. doi: 10.1016/j.neuroimage.2013.08.064

Gold, B. T., Johnson, N. F., and Powell, D. K. (2013). Lifelong bilingualism contributes to cognitive reserve against white matter integrity declines in aging. *Neuropsychologia* 51, 2841–2846. doi: 10.1016/j.neuropsychologia.2013.09.037

Grady, C. L., Luk, G., Craik, F. I. M., and Bialystok, E. (2015). Brain network activity in monolingual and bilingual older adults. *Neuropsychologia* 66, 170–181. doi: 10.1016/j.neuropsychologia.2014.10.042

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingual. Lang. Cogn.* 1, 67–81. doi: 10.1017/S1366728998000133

Green, D. W., and Abutalebi, J. (2008). Understanding the link between bilingual aphasia and language control. *J. Neurolinguist.* 21, 558–576. doi: 10.1016/j.jneuroling.2008.01.002

Green, D. W., and Abutalebi, J. (2013). Language control in bilinguals : the adaptive control hypothesis Language control in bilinguals : the adaptive control hypothesis. *J. Cogn. Psychol.* 25, 515–530. doi: 10.1080/20445911.2013.796377

Grogan, A., Jones, O. P., Ali, N., Crinion, J., Orabona, S., Mechias, M. L., et al. (2012). Structural correlates for lexical efficiency and number of languages in non-native speakers of English. *Neuropsychologia* 50, 1347–1352. doi: 10.1016/j.neuropsychologia.2012.02.019

Grundy, J. G., Anderson, J. A., and Bialystok, E. (2017). Bilinguals have more complex EEG brain signals in occipital regions than monolinguals. *NeuroImage* 159, 280–288. doi: 10.1016/j.neuroimage.2017.07.063

Hämäläinen, S., Sairanen, V., Leminen, A., and Lehtonen, M. (2017). Bilingualism modulates the white matter structure of language-related pathways. *Neuroimage* 152, 249–257. doi: 10.1016/j.neuroimage.2017.02.081

Han, Z., Ma, Y., Gong, G., He, Y., Caramazza, A., and Bi, Y. (2013). White matter structural connectivity underlying semantic processing: evidence from brain damaged patients. *Brain* 136, 2952–2965. doi: 10.1093/brain/awt205

Hernández, S. E., Suero, J., Barros, A., and González-Mora, J. L. (2016). Increased grey matter associated with long- term sahaja yoga meditation : a voxel-based morphometry study. *PLoS ONE* 11:e0150757. doi: 10.1371/journal.pone.0150757

Hervais-Adelman, A., Moser-Mercer, B., Michel, C. M., and Golestani, N. (2014). fMRI of simultaneous interpretation reveals the neural basis of extreme language control. *Cereb. Cortex* 25, 4727–4739. doi: 10.1093/cercor/bhu158

Hickok, G., and Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition* 92, 67–99. doi: 10.1016/j.cognition.2003.10.011

Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402. doi: 10.1038/nrn2113

Hopp, H. (2010). Ultimate attainment in L2 inflection: performance similarities between non-native and native speakers. *Lingua* 120, 901–931. doi: 10.1016/j.lingua.2009.06.004

Hopp, H. (2014). Working memory effects in the L2 processing of ambiguous relative clauses. *Lang. Acquis.* 21, 250–278. doi: 10.1080/10489223.2014.892943

Hosoda, C., Tanaka, K., Nariai, T., Honda, M., and Hanakawa, T. (2013). Dynamic neural network reorganization associated with second language vocabulary acquisition: a multimodal imaging study. *J. Neurosci.* 33, 13663–13672. doi: 10.1523/JNEUROSCI.0410-13.2013

Johnson, J. S., and Newport, E. L. (1991). Critical period effects on universal properties of language: the status of subjacency in the acquisition of a second language. *Cognition* 39, 215–258. doi: 10.1016/0010-0277(91)90054-8

Kennedy, K., and Raz, N. (2009). Aging white matter and cognition: differential effects of regional variations in diffusion properties on memory, executive functions, and speed. *Neuropsychologia* 47, 916–927. doi: 10.1016/j.neuropsychologia.2009.01.001

Kousaie, S., Chai, X. J., Sander, K. M., and Klein, D. (2017). Simultaneous learning of two languages from birth positively impacts intrinsic functional connectivity and cognitive control. *Brain Cogn.* 117, 49–56. doi: 10.1016/j.bandc.2017.06.003

Kroll, J. F., Bobb, S. C., and Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: arguments against a fixed locus of language selection in bilingual speech. *Bilingualism* 9, 119–135. doi: 10.1017/S1366728906002483

Kuhl, P. K., Stevenson, J., Corrigan, N. M., van den Bosch, J. J., Can, D. D., and Richards, T. (2016). Neuroimaging of the bilingual brain: structural brain correlates of listening and speaking in a second language. *Brain Lang.* 162, 1–9. doi: 10.1016/j.bandl.2016.07.004

Kunimatsu, A., Aoki, S., Masutani, Y., Abe, O., Hayashi, N., Mori, H., et al. (2004). The optimal trackability threshold of fractional anisotropy for diffusion tensor tractography of the corticospinal tract. *Magn. Reson. Med. Sci.* 3, 11–17. doi: 10.2463/mrms.3.11

Leclercq, D., Duffau, H., Delmaire, C., Capelle, L., Gatignol, P., Ducros, M., et al. (2010). Comparison of diffusion tensor imaging tractography of language tracts and intraoperative subcortical stimulations. *J. Neurosurg.* 112, 503–511. doi: 10.3171/2009.8.JNS09558

Li, P., Legault, J., and Litcofsky, K. A. (2014). Neuroplasticity as a function of second language learning: anatomical changes in the Human brain. *Cortex* 58, 301–324. doi: 10.1016/j.cortex.2014.05.001

Linck, J. A, Kroll, J. F., and Sunderman, G. (2009). Losing access to the native language while being immersed in a second language. Evidence for the role of inhibition in second language learning. *Psychol. Sci.* 20, 1507–1515. doi: 10.1111/j.1467-9280.2009.02480.x

Linck, J. A., Osthus, P., Koeth, J. T., and Bunting, M. F. (2014). Working memory and second language comprehension and production: a meta-analysis. *Psychon. Bull. Rev.* 21, 861–883. doi: 10.3758/s13423-013-0565-2

Long, M. H. (1990). Maturational constraints on language development. *Stud. Second Lan. Acquis.* 12, 251–285. doi: 10.1017/S0272263100009165

López-Barroso, D., Catani, M., Ripollés, P., Dell'Acqua, F., Rodríguez-Fornells, A., and de Diego-Balaguer, R. (2013). Word learning is mediated by the left arcuate fasciculus. *Proc. Natl. Acad. Sci. U.S.A.* 110, 13168–13173. doi: 10.1073/pnas.1301696110

Luk, G., and Pliatsikas, C. (2016). Converging diversity to unity: commentary on The neuroanatomy of bilingualism. *Lang. Cogn. Neurosci.* 31, 349–352. doi: 10.1080/23273798.2015.1119289

Luk, G., Bialystok, E., Craik, F. I. M., and Grady, C. L. (2011). Lifelong bilingualism maintains white matter integrity in older adults. *J. Neurosci.* 31, 16808–16813. doi: 10.1523/JNEUROSCI.4563-11.2011

MacWhinney, B. (2005). "A unified model of language acquisition," *Handbook of Bilingualism: Psycholinguistic Approaches*, eds J. F. Kroll and A. M. B. de groot (New York, NY: Oxford University Press), 49–67.

Maguire, E. A, Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S., et al. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *Proc. Natl. Acad. Sci. U.S.A.* 97, 4398–4403. doi: 10.1073/pnas.070039597

Mamiya, P. C., Richards, T. L., Coe, B. P., Eichler, E. E., and Kuhl, P. K. (2016). Brain white matter structure and COMT gene are linked to second-language learning in adults. *Proc. Natl. Acad. Sci. U.S.A.* 113, 7249–7254. doi: 10.1073/pnas.1606602113

Martino, J., Brogna, C., Robles, S. G., Vergani, F., and Duffau, H. (2010). Anatomic dissection of the inferior fronto-occipital fasciculus revisited in the lights of brain stimulation data. *Cortex* 46, 691–699. doi: 10.1016/j.cortex.2009.07.015

McDonald, J. L. (2000). Grammaticality judgments in a second language: influences of age of acquisition and native language. *Appl. Psycholinguist.* 21, 395–423. doi: 10.1017/S0142716400003064

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.* 2, 123–150. doi: 10.1111/j.1467-9922.2010.00604.x

McMurray, B., Samelson, V. M., Lee, S. H., and Tomblin, J. B. (2010). Individual differences in online spoken word recognition: implications for SLI. *Cogn. Psychol.* 60, 1–39. doi: 10.1016/j.cogpsych.2009.06.003

Mechelli, A., Crinion, J., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiack, R., et al. (2004). Structural plasticity in the bilingual brain. *Nature* 431:757. doi: 10.1038/431757a

Ministry of Education Culture and Sport of Spain (2006). *Diploma de Español como Lengua Extranjera (Diploma of Spanish as a Second Language Examination)*. Available online at: http://www.dele.org/

Mirman, D., Zhang, Y., Wang, Z., Coslett, H. B., and Schwartz, M. F. (2015). The ins and outs of meaning: behavioral and neuroanatomical dissociation of semantically-driven word retrieval and multimodal semantic recognition in aphasia. *Neuropsychologia* 76, 208–219. doi: 10.1016/j.neuropsychologia.2015.02.014

Mohades, S. G., Van Schuerbeek, P., Rosseel, Y., Van De Craen, P., Luypaert, R., and Baeken, C. (2015). White-matter development is different in bilingual and monolingual children: a longitudinal DTI study. *PLoS ONE* 10:e0117968. doi: 10.1371/journal.pone.0117968

Mohades, S., Struys, E., Van Schuerbeek, P., Mondt, K., Van De Craen, P., and Luypaert, R. (2012). DTI reveals structural differences in white matter tracts between bilingual and monolingual children. *Brain Res.* 1435, 72–80. doi: 10.1016/j.brainres.2011.12.005

Nichols, E. S., and Joanisse, M. F. (2016). Functional activity and white matter microstructure reveal the independent effects of age of acquisition and proficiency on second-language learning. *Neuroimage* 143, 15–25. doi: 10.1016/j.neuroimage.2016.08.053

Olsen, R. K., Pangelinan, M. M., Bogulski, C., Chakravarty, M. M., Luk, G., Grady, C. L., et al. (2015). The effect of lifelong bilingualism on regional grey and white matter volume. *Brain Res.* 1612, 128–139. doi: 10.1016/j.brainres.2015.02.034

Papagno, C. (2011). Naming and the role of the uncinate fasciculus in language function. *Curr. Neurol. Neurosci. Rep.* 11, 553. doi: 10.1007/s11910-011-0219-6

Perani, D., and Abutalebi, J. (2005). Neural basis of first and second language processing. *Curr. Opin. Neurobiol.* 15, 202–206. doi: 10.1016/j.conb.2005.03.007

Perani, D., and Abutalebi, J. (2015). Bilingualism, dementia, cognitive and neural reserve. *Curr. Opin. Neurol.* 28, 618–625. doi: 10.1097/WCO.0000000000000267

Perani, D., Dehaene, S., Grassi, F., Cohen, L., Cappa, S. F., Dupoux, E., et al. (1996). Brain processing of native and foreign languages. *Neuroreport* 7, 2439–2444. doi: 10.1097/00001756-199611040-00007

Perani, D., Paulesu, E., Galles, N. S., Dupoux, E., Dehaene, S., Bettinardi, V., et al. (1998). The bilingual brain. Proficiency and age of acquisition of the second language. *Brain* 121, 1841–1852. doi: 10.1093/brain/121.10.1841

Pivneva, I., Palmer, C., and Titone, D. (2012). Inhibitory control and L2 proficiency modulate bilingual language production: evidence from spontaneous monologue and dialogue speech. *Front. Psychol.* 3:57. doi: 10.3389/fpsyg.2012.00057

Pliatsikas, C., and Chondrogianni, V. (2015). Editorial: learning a non-native language in a naturalistic environment: insights from behavioral and neuroimaging research. *Front. Psychol.* 6:1009. doi: 10.3389/978-2-88919-639-5

Pliatsikas, C., DeLuca, V., Moschopoulou, E., and Saddy, J. D. (2017). Immersive bilingualism reshapes the core of the brain. *Brain Struct. Funct.* 222, 1785–1795. doi: 10.1007/s00429-016-1307-9

Pliatsikas, C., Moschopoulou, E., and Saddy, J. D. (2015). The effects of bilingualism on grey and white matter structure. *Proc. Natl. Acad. Sci. U.S.A.* 112, 1334–1337. doi: 10.1073/pnas.1414183112

Qi, Z., Han, M., Garel, K., San, E., and Gabrieli, J. D. E. (2015). White-matter structure in the right hemisphere predicts Mandarin Chinese learning success. *J. Neurolinguist.* 33, 14–28. doi: 10.1016/j.jneuroling.2014.08.004

Rahmani, F., Sobhani, S., and Aarabi, M. H. (2017). Sequential language learning and language immersion in bilingualism: diffusion MRI connectometry reveals microstructural evidence. *Exp. Brain Res.* 235, 2935–2945. doi: 10.1007/s00221-017-5029-x

Ripollés, P., Marco-Pallarés, J., Hielscher, U., Mestres-Missé, A., Tempelmann, C., Heinze, H. J., et al. (2014). The role of reward in word learning and its implications for language acquisition. *Curr. Biol.* 24, 2606–2611. doi: 10.1016/j.cub.2014.09.044

Roelofs, A. (2014). A dorsal-pathway account of aphasic language production: the WEAVER++/ARC model. *Cortex* 59, 33–48. doi: 10.1016/j.cortex.2014.07.001

Rosenzweig, M. R., Krech, D., Bennett, E. L., Diamond, M. C. (1962). Effects of environmental complexity and training on brain chemistry and anatomy: a replication and extension. *J. Comp. Physiol. Psychol.* 55, 429–437. doi: 10.1037/h0041137

Rossi, E., Kroll, J. F., and Dussias, P. E. (2014). Clitic pronouns reveal the time course of processing gender and number in a second language. *Neuropsychologia* 62, 11–25. doi: 10.1016/j.neuropsychologia.2014.07.002

Rueckert, D., Sonoda, L. I., Hayes, C., Hill, D. L., Leach, M. O., and Hawkes, D. J. (1999). Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Trans. Med. Imaging* 18, 712–721. doi: 10.1109/42.796284

Sabourin, L., and Stowe, L. A. (2008). Second language processing: when are first and second languages processed similarly? *Second Lang. Res.* 24, 397–430. doi: 10.1177/0267658308090186

Sabourin, L., Stowe, L. A., and de Haan, G. J. (2006). Transfer effects in learning a second language grammatical gender system. *Second Lang. Res.* 22, 1–29. doi: 10.1191/0267658306sr259oa

Schlegel, A. A., Rudelson, J. J., and Tse, P. U. (2012). White matter structure changes as adults learn a second language. *J. Cogn. Neurosci.* 24, 1664–1670. doi: 10.1162/jocn_a_00240

Schnur, T. T., Schwartz, M. F., Brecher, A., and Hodgson, C. (2006). Semantic interference during blocked-cyclic naming: evidence from aphasia. *J. Mem. Lang.* 54, 199–227. doi: 10.1016/j.jml.2005.10.002

Skare, S., Forsman, L., and Bengtsson, S. L. (2005). Extensive piano practicing has regionally specific effects on white matter development. *Nat. Neurosci.* 8, 1148–1151.

Smith, S. M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T. E., MacKay, C. E., et al. (2006). Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage* 31, 1487–1505. doi: 10.1016/j.neuroimage.2006.02.024

Smith, S. M. (2002). Fast robust automated brain extraction. *Hum. Brain Mapp.* 17, 143–155. doi: 10.1002/hbm.10062

Smith, S., Jenkinson, M., Woolrich, M., Beckmann, C., Behrens, T., Johansen-Berg, H., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23, 208–219. doi: 10.1016/j.neuroimage.2004.07.051

Stein, M., Federspiel, A., Koenig, T., Wirth, M., Strik, W., Wiest, R., et al. (2012). Structural plasticity in the language system related to increased second language proficiency. *Cortex* 48, 458–465. doi: 10.1016/j.cortex.2010.10.007

Stein, M., Winkler, C., Kaiser, A., and Dierks, T. (2014). Structural brain changes related to bilingualism: does immersion make a difference? *Front. Psychol.* 5:1116. doi: 10.3389/fpsyg.2014.01116

Tan, L. H., Chen, L., Yip, V., Chan, A. H., Yang, J., Gao, J. H., et al. (2011). Activity levels in the left hemisphere caudate–fusiform circuit predict how well a second language will be learned. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2540–2544. doi: 10.1073/pnas.0909623108

Teichmann, M., Rosso, C., Martini, J. B., Bloch, I., Brugières, P., Duffau, H., et al. (2015). A cortical–subcortical syntax pathway linking Broca's area and the striatum. *Hum. Brain Mapp.* 36, 2270–2283. doi: 10.1002/hbm.22769

Ullman, M. T. (2016). "The declarative/procedural model: a neurobiological model of language learning, knowledge, and use," in *Neurobiology of Language*, eds G. Hickok and S. L. Small (Cambridge, MA: Academic Press), 953–968. doi: 10.1016/B978-0-12-407794-2.00076-6

Wakana, S., Caprihan, A., Panzenboeck, M. M., Fallon, J. H., Perry, M., Gollub, R. L., et al. (2007). Reproducibility of quantitative tractography methods applied to cerebral white matter. *Neuroimage* 36, 630–644. doi: 10.1016/j.neuroimage.2007.02.049

Wartenburger, I., Heekeren, H. R., Abutalebi, J., Cappa, S. F., Villringer, A., and Perani, D. (2003). Early setting of grammatical processing in the bilingual brain. *Neuron* 37, 159–170. doi: 10.1016/S0896-6273(02)01150-9

Weber-Fox, C. M., and Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *J. Cogn. Neurosci.* 8, 231–256. doi: 10.1162/jocn.1996.8.3.231

Xiang, H., van Leeuwen, T. M., Dediu, D., Roberts, L., Norris, D. G., and Hagoort, P. (2015). L2-proficiency-dependent laterality shift in structural connectivity of brain language pathways. *Brain Connect.* 5, 349–361. doi: 10.1089/brain.2013.0199

Xu, M., Baldauf, D., Chang, C. Q., Desimone, R., and Tan, L. H. (2017). Distinct distributed patterns of neural activity are associated with two languages in the bilingual brain. *Science Adv.* 3:e1603309. doi: 10.1126/sciadv.1603309

Zou, L., Abutalebi, J., Zinszer, B., Yan, X., Shu, H., Peng, D., et al. (2012a). Second language experience modulates functional brain network for the native language production in bimodal bilinguals. *Neuroimage* 62, 1367–1375. doi: 10.1016/j.neuroimage.2012.05.062

Zou, L., Ding, G., Abutalebi, J., Shu, H., and Peng, D. (2012b). Structural plasticity of the left caudate in bimodal bilinguals. *Cortex* 48, 1197–1206. doi: 10.1016/j.cortex.2011.05.022

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 A: LANGUAGE HISTORY QUESTIONNAIRE

Language History Questionnaire

This questionnaire is designed to give us a better understanding of your experience with other languages. We ask that you be as accurate as thorough as possible when answering the following questions.

1. Gender

   ❏ Male
   ❏ Female

2. Age:

3. Do you have any known visual or hearing problems (corrected or uncorrected)?

   ❏ Yes
   ❏ No

4. Native Country

   ❏ United States
   ❏ Other ————————————————

5. Native Language

   ❏ English
   ❏ Other ————————————————

6. Language(s) spoken at home (Please check all that apply).

   ❏ English
   ❏ Spanish
   ❏ German
   ❏ Other [Please explain: ————————————— –

If ENGLISH is your Native Language, please RATE yourself:
***If English is NOT your Native Language, please contact Experimenter for further instructions.

7. Please rate your English reading proficiency. (1=not literate and 10 = very literate)

   ❏ 1   ❏ 6
   ❏ 2   ❏ 7
   ❏ 3   ❏ 8
   ❏ 4   ❏ 9
   ❏ 5   ❏ 10

8. Please rate your English writing proficiency. (1=not literate and 10=very literate)

   ❏ 1   ❏ 6
   ❏ 2   ❏ 7
   ❏ 3   ❏ 8
   ❏ 4   ❏ 9
   ❏ 5   ❏ 10

9. Please rate your English speaking ability. (1=not fluent and 10=very fluent)

   ❏ 1   ❏ 6
   ❏ 2   ❏ 7
   ❏ 3   ❏ 8
   ❏ 4   ❏ 9
   ❏ 5   ❏ 10

9. Please rate your English speech comprehension ability. (1=unable to understand conversation and 10=perfectly able to understand)

   ❏ 1   ❏ 6
   ❏ 2   ❏ 7
   ❏ 3   ❏ 8
   ❏ 4   ❏ 9
   ❏ 5   ❏ 10

The next section of the questionnaire deals with your second language learning experience.

11. Have you studied any second language?

    ❏ No →**If NO, please go to question #19**
    ❏ Yes

    If YES, where and when? Please check all that apply and indicate length of study.

    ❏ Home Language: ————————————
    ❏ Since Age ( )

Elementary School Language: ————————————-

❏ ( ) year(s)

Middle School Language: — ————————————

❏ ( ) year(s)

High School Language: — ————————————

❏ 1 year
❏ 2 years
❏ 3 years

College Language:

❏ Have not studied a second language in college
❏ 1-2 semesters
❏ 3-4 semesters
❏ 5-6 semesters
❏ 8+ semesters

12. If you are taking/have taken any second language at college, please answer the following question. Are you: (Please check all that apply.)

    ❏ Taking a second language for a requirement but interested in being a major or minor.
    ❏ A second language minor
    ❏ A second language major
    ❏ A second language graduate student
    ❏ Other [please explain ————————————-]

13. Have you studied / lived abroad?

❏ Yes
❏ No

If Yes, where and when did you study, for how long, and what language did you speak?

Country	Approx. dates	Length of Stay	Language

The next section asks you to rate your skills in your primary second language.

14. Please rate your second language reading proficiency. (1=not literate and 10=very literate)

   ❏ 1   ❏ 6
   ❏ 2   ❏ 7
   ❏ 3   ❏ 8
   ❏ 4   ❏ 9
   ❏ 5   ❏ 10

15. Please rate your second language writing proficiency. (1=not literate and 10=very literate)

   ❏ 1   ❏ 6
   ❏ 2   ❏ 7
   ❏ 3   ❏ 8
   ❏ 4   ❏ 9
   ❏ 5   ❏ 10

16. Please rate your second language speaking ability. (1=not fluent and 10=very fluent)

   ❏ 1   ❏ 6
   ❏ 2   ❏ 7
   ❏ 3   ❏ 8
   ❏ 4   ❏ 9
   ❏ 5   ❏ 10

17. Please rate your second language speech comprehension ability. (1=unable to understand conversation and 10=perfectly able to understand)

   ❏ 1   ❏ 6
   ❏ 2   ❏ 7
   ❏ 3   ❏ 8
   ❏ 4   ❏ 9
   ❏ 5   ❏ 10

18. In my second language class es I get:

❏ Mostly A's
❏ Mostly A's and B's
❏ Mostly B's
❏ Mostly B's and C's
❏ Mostly C's

19. If you speak or have studied more than one second language, please explain about your additional language experience (i.e., years, level of proficiency, etc.)

Thank you for your participation

Check for updates

Differential Signatures of Second Language Syntactic Performance and Age on the Structural Properties of the Left Dorsal Pathway

Kayako Yamamoto[1,2] and Kuniyoshi L. Sakai[1,3]*

[1] Department of Basic Science, Graduate School of Arts and Sciences, The University of Tokyo, Tokyo, Japan, [2] Japan Society for the Promotion of Science, Saitama, Japan, [3] Japan Agency for Medical Research and Development – Core Research for Evolutional Science and Technology, Tokyo, Japan

In adult second language (L2) acquisition, individual differences are considerably large even among people with similar experiences. The neural mechanisms underlying this variability would include structural plasticity of language-related pathways. To elucidate such neuroplasticity, we focused on the transitional period of adolescence, which is associated with certain plasticity toward maturation following the sensitive period of language acquisition ($\leq$12 years old). The adolescent brain would thus be influenced by age-dependent factors, as well as performances in L2. Here, we examined individual differences in L2 performances controlling the duration of experience to reveal the differential signatures of performances and age on the plasticity of structural properties in major language-related pathways. We recruited Japanese students at two ages, i.e., junior (age: 13–14) and senior (age: 16–17) high-school students, all of whom started to expose to English at age 12 or 13. We divided them into subgroups, so that either L2 performance [Junior (High)/Senior (Low)] or age [Senior (Low)/Senior (High)] was matched in group comparisons; the duration of L2 experience was also controlled between the Senior (Low) and Senior (High) groups. We then examined the thickness and fractional anisotropy (FA) of the dorsal and ventral pathways, i.e., the arcuate fasciculus (Arcuate) and inferior fronto-occipital fasciculus (IFOF), respectively, using semi-automatic methods for selecting regions without branches. Regarding FA in the left Arcuate, the Senior (High) group showed significantly higher FA than the other two groups, indicating performance-related group differences. Further, FA in the left Arcuate was selectively correlated with the accuracy of a syntactic task. Regarding the thickness of the left Arcuate, the Senior (High) and Senior (Low) groups showed significantly larger thickness than the Junior (High) group, indicating age-related group differences. These differential performance-related and age-related signatures were evident on the left Arcuate alone, in contrast to the right Arcuate that showed only mild differences in thickness, and to the bilateral IFOF that lacked either signature. Our results suggest that the left dorsal pathway continued to develop to adolescence, and that performance differences in a syntactic task can be predicted by its FA, independent of age and the duration of experience.

Keywords: diffusion MRI, white matter, dorsal and ventral pathways, language acquisition, syntax

INTRODUCTION

Second language (L2) acquisition shows considerably large individual differences, especially when the L2 is acquired in adulthood. Even among people with similar L2 experiences (e.g., taking the same classes or lessons on a foreign language), some improve their L2 performances in a relatively short period of time, while others do not improve as well. This is in marked contrast to first language (L1) acquisition, in which linguistic abilities are similar among individuals despite highly variable experiences. However, individual differences in L1 do emerge when effortful language use is imposed. By using verb-, rhyme-, and opposite word generation tasks in L1 for 9-year-old children and adult participants, previous studies have reported that distinct regions in the left frontal and occipital cortices showed age-related or performance-related activations (Schlaggar et al., 2002; Brown et al., 2005). These studies have indicated the importance of employing subgroups, in which either age or performance was matched, thereby fixing one of the two factors in group comparisons. Individual differences would also be revealed when people learn to read or write even with their own languages through educational training of specific skills. A previous study has reported that the structural property of a white matter pathway connecting the left temporal and parietal language areas may have plasticity associated with literacy experience even for adults, by comparing three groups of illiterates, ex-illiterates who learned to read during adulthood, and literates (Thiebaut de Schotten et al., 2014). Long and intensive experiences are required also in L2 acquisition, which may be supported by the structural plasticity of language-related pathways. To elucidate such neuroplasticity, it is of interest to focus on the transitional period of adolescence, which occurs after the sensitive period of language acquisition (≤12 years old). Indeed, adolescence has been suggested to involve certain plasticity toward maturation (Fuhrmann et al., 2015). The total cerebral volume has been reported to show a gentle inverted-U trend, with a peak age different among genders (girls: 10.5 years, boys: 14.5 years) (Lenroot et al., 2007). While the gray matter volume gradually decreases in adolescence, after peaking at later years in higher order association areas (Gogtay et al., 2004; Lenroot et al., 2007), the white matter continues to increase until the twenties or later with some regional differences (Westlye et al., 2010). Moreover, developments in white-matter pathways are accompanied by large individual differences. A previous longitudinal study showed that some show increases in volume, while others show decreases, even at similar ages (Lebel and Beaulieu, 2011). The white matter in the adolescent brain would be influenced by multiple factors depending on the participants' ages (e.g., biological maturation), as well as on the attainment of cognitive/motor abilities. More specifically, in major language-related pathways, the independent factors of age and performances after intensive L2 experiences may be reflected in the plasticity of different structural properties.

Given the large individual differences in L2, many complicated issues should be tackled in examining the neural plasticity related to L2 acquisition. In addition to participants' current age, the age of first exposure (AOE) and the duration of exposure (DOE) represent other factors in L2 acquisition (Li et al., 2014). In the present study, we controlled AOE in order to examine any group differences related to the current ages of participants (hereafter, age-related group differences), and therefore recruited students at two ages, i.e., junior (age: 13–14) and senior (age: 16–17) high-school students. Their AOE to English was 12 or 13, and they attended the same school where English classes were based on the curriculum guidelines determined by MEXT (Ministry of Education, Culture, Sports, Science and Technology). The temporal factor of L2 experience, as represented by the DOE, was also controlled and eliminated from the performances for the students at the same age. Even among such students, large individual variations in L2 performance were observable. Given that Japanese students tend to make similar mistakes in English, such as applying the null-subject (pro-drop) allowed in Japanese, we examined participants' syntactic abilities in English by a syntactic error-detection task (Syn) that we previously developed with high school teachers (Sakai et al., 2009). Orthographic knowledge in English was also examined by using a spelling error-detection task (Spe) with basically the same sets of sentences. Based on the task performances, the junior and senior students were separately divided into subgroups. As a result, the task performances of the Junior (High) and the Senior (Low) groups matched, while those of the Senior (High) group were significantly better. We then compared these three groups, with either age or L2 performance being fixed. After the group division, we identified the dorsal and ventral language-related pathways, two major routes that combine multiple language areas (Hickok and Poeppel, 2007; Friederici, 2011). The dorsal and ventral pathways correspond to the fronto-temporal segment of the arcuate fasciculus (Arcuate) and the inferior fronto-occipital fasciculus (IFOF), respectively. We identified these pathways using our previously established semi-automatic methods of defining seeds for tractography in a diffusion magnetic resonance imaging (diffusion MRI) analysis (Yamamoto and Sakai, 2016). We focused on two distinct structural properties, i.e., the thickness (or the volume of a tract) and fractional anisotropy (FA—the degree of myelination and/or fiber organization), to examine how age and L2 performance are reflected in these properties of the dorsal and ventral pathways in each hemisphere.

Several diffusion MRI studies have suggested that the dorsal pathways mature later than the ventral pathways in both infants and children. A diffusion MRI study with tractography has reported that one of the two dorsal pathways, the one connecting the inferior frontal gyrus (IFG) and the temporal cortex, was not trackable in 2-day-old newborns; in contrast, the IFOF of the ventral pathways, as well as the other dorsal pathway connecting the premotor cortex and the temporal cortex, was already present (Perani et al., 2011). In a study with 7-year-old children, both the dorsal and ventral pathways were trackable, but FA of these pathways was significantly lower than that of adults (Brauer et al., 2013), suggesting that the language-related pathways were not fully mature at this age. Regarding the adolescent ages, it has been suggested that the dorsal pathway was still under development, and that the ventral pathway showed less prominent development (Lebel and Beaulieu, 2011). FA in the dorsal pathway has been

reported to increase during adolescence, mainly associated with increases in parallel diffusivity, which may reflect the increases in axon diameters (Giorgio et al., 2010). FA in the language-related pathways may increase with linguistic experiences, as previous longitudinal studies reported that FA of the left ventral pathway of bilinguals was higher than that of monolinguals, possibly reflecting differences in semantic aspects (Mohades et al., 2015). However, it remains less clear how these structural developments can be correlated with behavioral measurements. The relatively slow development of the dorsal pathway suggests the interesting possibility that this pathway may be more strongly influenced by language experiences. As suggested by a native language neural commitment theory (Kuhl, 2004), neural networks develop under the coding of L1 inputs from early ages, further influencing the acquisition of L2 in later ages. Indeed, L2 abilities have been reported to be correlated with L1 abilities, for instance in reading proficiency (van Gelderen et al., 2007). Given that improvements, as well as individual differences, in L1 during adolescence become smaller than those in L2, examining the relationships between L2 acquisition and the structural properties would be of worth in understanding the neural plasticity of language-related pathways in adolescence.

We have performed the following functional and structural studies on adolescent students. In our previous functional magnetic resonance imaging (fMRI) study, we identified cortical regions involved in syntactic processing using the same English tasks as in the present study, and showed a positive correlation between the individual activations of the left IFG and the performance accuracy of the Syn task (Sakai et al., 2009). Our voxel-based morphometry (VBM) study further clarified that individual leftward lateralization of a single region in the IFG also showed a correlation with the accuracy of Syn (Nauchi and Sakai, 2009). Moreover, in our recent diffusion MRI study we reported that individual FA in the left Arcuate was correlated with the accuracy of Syn (Yamamoto and Sakai, 2016), suggesting the importance of the left dorsal network including the IFG and the Arcuate in syntactic processing. More specifically, among the Arcuate and IFOF in both hemispheres, we revealed that FA in the left Arcuate alone was positively correlated with the performance, i.e., accuracy of the Syn task, but not with that of Spe or with L1 verbal fluency. Further, within monozygotic twin pairs, neither the accuracy of Syn nor FA in the left Arcuate were significantly correlated between the twins, in spite of the high inter-twin correlation for the thickness of the left Arcuate. Given these results, syntactic abilities in L2 and FA in the left Arcuate may thus have been sensitive to non-shared environmental factors by which the twins were individually affected, while the thickness was dependent on shared genetic/environmental factors. Based on these points, we made two further hypotheses. First, FA in the left Arcuate should show performance-related group differences that would be observable even among the students with the same DOE, because larger variances within the monozygotic twin pairs were observed for FA than for the thickness. Second, the thickness of the left Arcuate should show age-related group differences, as a similar thickness was observed for monozygotic twin pairs. We tested these hypotheses using semi-automatic methods, which we developed previously for

defining seeds in tractography and selecting regions of interest (ROIs) (Yamamoto and Sakai, 2016). In this previous study, we reported that the thickness of the left Arcuate, averaged in a one-dimensional ROI, was clearly larger than that of the right Arcuate; such laterality was evident neither for the thickness of the IFOF, nor for FA in the Arcuate or IFOF. In the present study, recruiting a larger number of adolescent participants, we confirmed that the leftward laterality of the Arcuate was consistent among the three groups. We further examined the correlation with individual accuracy of Syn/Spe to clarify which particular aspects of L2 abilities were related to the structural properties that showed any performance-related group differences. Our results should thus help to elucidate the neuroanatomical basis of language acquisition after the sensitive period.

MATERIALS AND METHODS

Participants

Junior high-school students (age: 13–14) and senior high-school students (age: 16-17) were recruited from the Secondary School attached to the Faculty of Education, the University of Tokyo. Twin pairs were included among the participants; this school has performed educational as well as twin research. The following accumulative inclusion criteria were employed in the present study: (i) right-handedness, i.e., a positive laterality quotient (LQ), which was assessed using the Edinburgh inventory (Oldfield, 1971), (ii) with neither hearing/visional problems nor history of neurological/psychiatric diseases, (iii) native Japanese speakers who started to acquire English in formal education at the age of 12 or 13 (this condition was met by most of the students in this school), and (iv) reaction times (RTs) for Spe within the presentation time (6400 ms) for more than 90% of the trials including incorrect responses. To avoid including monozygotic/dizygotic twin pairs with potentially similar characteristics, the one with the higher score in Spe was analyzed for each pair who met these criteria. As regards the first criterion, 8 junior and 11 senior students, who showed a negative LQ or reported a potential history of left-handedness, were dropped. As a result, the population showed relatively strong right-handedness (LQ > 35). One participant each for the second and fourth criterion, and two participants for the third criterion, were dropped.

For the senior high-school students, we employed three additional criteria: (v) an accuracy of Spe higher than 65% (i.e., "mean − 1.5 SD"), (vi) no worse outliers in each task (i.e., higher than the "mean − 2 SD" for the accuracy, and shorter than the "mean + 2 SD" for the RTs), and (vii) shorter RTs for easier Spe than for Syn. Given that most Japanese students learn an alphabetic writing system at the age of 12–13, and that the senior high-school students had been studying English for about 4 years, the fifth criterion was necessary to exclude the potential effects of poor reading abilities and precisely assess individual syntactic abilities; five senior high-school students were dropped for this reason. Two students were dropped because they did not meet the sixth criterion, and two more students because they did not

meet the seventh criterion. As a result, we enrolled a total of 39 junior and 38 senior high-school students.

We divided the junior high-school students into two groups: a group of 14 students [the Junior (High) group] who scored higher than 65% in Spe (this was identical to the fifth criterion employed for senior high-school students), and a group of 25 students [the Junior (Low) group] with scores lower than 65% in Spe, which were too low to assess their L2 abilities and related structures further. In regard to the senior high-school students, we first divided the 38 students into two groups using K-means cluster analysis (R software)[1] on the accuracy of Syn and Spe tasks. This analysis yielded a group of 15 participants with higher L2 abilities [the Senior (High) group], as well as a group of 23 participants with lower L2 abilities. Because the performances of the group with lower L2 abilities did not match those of the Junior (High) group for the accuracy and RTs of Spe (one-sided t-tests, $P < 0.05$), we further performed a hierarchical cluster analysis using Ward's method in the R software[1] based on the accuracy and RTs of Spe for the 23 students with lower L2 abilities. As a result, we divided these students into two groups: a group of 5 students with the higher accuracy and shorter RTs of Spe [the Senior (Middle) group], and a group of 18 other students [the Senior (Low) group]. The accuracy/RTs of both tasks for the Senior (Low) group matched those for the Junior (High) group (one-sided t-tests, $P > 0.05$). Because our purpose was to examine the differences between the performance-matched groups with different ages, as well as between age-matched groups with different L2 abilities, we focused on the Junior (High), Senior (Low), and Senior (High) groups, dropping the other groups. All participants in the Senior (High) group, as well as seven participants in the Senior (Low) group, were included in our previous study (Yamamoto and Sakai, 2016), in which we analyzed participants with an accuracy of Spe ≥ 80% (Experiment I); moreover, these former participants were the subset of the same 38 senior high-school students mentioned above. Participants in the Junior (High) group were newly recruited for the present study. All participants in the three groups were scanned with the same diffusion MRI protocol. Demographic details of participants in these three groups, such

---

[1]https://www.r-project.org/

as age, AOE, DOE, and LQ, are shown in **Table 1**. We obtained written informed consent for this research from all the students and their guardians. This study was approved by the Institutional review board of the University of Tokyo (Komaba) and by the Secondary School.

Stimuli and Tasks

All the students performed English syntactic (Syn) and spelling (Spe) error-detection tasks (50 trials each for Syn and Spe tasks), as well as a verbal fluency task in Japanese (for three initial letters in hiragana). The tasks were the same as those we employed in our previous studies (Nauchi and Sakai, 2009; Sakai et al., 2009; Yamamoto and Sakai, 2016); the stimuli for Syn are listed in Sakai et al. (2009). In the Syn task, we tested argument structures and related syntactic knowledge, which are difficult for L2 learners to acquire. For instance, as objects of transitive verbs are freely omitted in Japanese as well as other languages, Japanese students tend to accept incorrect (*) English sentences, such as "*Do you often meet Mary? – *Yes, I often meet*" [see Sakai et al. (2009) for further information on other types of syntactic errors]. Because ungrammatical sentences in the Syn task cannot be judged as incorrect by semantic information alone (e.g., by internally translating English sentences into Japanese), the accuracy of the Syn task reflected individual syntactic abilities appropriately.

Magnetic Resonance Image Acquisition and Data Analyses

Magnetic resonance images were acquired with the same protocols and parameters as in our previous study (Yamamoto and Sakai, 2016). With the same semi-automatic procedures described in this previous study, we identified the Arcuate and IFOF in each individual hemisphere. In the present study with group comparisons, all the tracts were normalized and analyzed after the tracking in individual brains; in our previous study, tracts were analyzed in the native space to focus on individual variabilities of structural indices (Yamamoto and Sakai, 2016). Using the affine and non-linear transformation with FLIRT and FNIRT, all the tracts were spatially normalized to the Montreal Neurological Institute (MNI) space, and were binarized with the

---

TABLE 1 | Demographic details of the participants in each group.

		N	Age	AOE	DOE	LQ	vf
Junior (High)	Female	7	14 ± 0.5	12.7 ± 0.3	1.5 ± 0.3	82 ± 20	24 ± 3.7
	Male	7	14 ± 0.3	12.6 ± 0.3	1.4 ± 0.2	86 ± 13	20 ± 2.6
	Total	14	14 ± 0.4	12.6 ± 0.3	1.5 ± 0.3	84 ± 16	22 ± 3.6
Senior (Low)	Female	10	17 ± 0.3	12.6 ± 0.3	4.4 ± 0.2	82 ± 21	24 ± 5.6
	Male	8	17 ± 0.3	12.7 ± 0.4	4.4 ± 0.2	92 ± 16	21 ± 5.0
	Total	18	17 ± 0.3	12.7 ± 0.3	4.4 ± 0.2	86 ± 19	23 ± 5.4
Senior (High)	Female	9	17 ± 0.4	12.5 ± 0.4	4.5 ± 0.2	82 ± 16	26 ± 6.5
	Male	6	17 ± 0.3	12.6 ± 0.3	4.5 ± 0.2	83 ± 16	29 ± 5.3
	Total	15	17 ± 0.4	12.6 ± 0.3	4.5 ± 0.2	83 ± 15	27 ± 6.0

Data are shown as the means ± SD for the number of participants (N), age, age of first exposure (AOE) to English, duration of exposure (DOE) to English, laterality quotient (LQ), and the number of words produced in the verbal fluency task (vf).

fslmaths function of FSL software. While FLIRT and FNIRT are conventional volume-based registration algorithms using voxel-level information, a method of combining volume-based and surface-based registration (Zöllei et al., 2010), as well as an improved voxel-based registration (Schwarz et al., 2014), has been proposed for better alignment. We overlaid the binarized tracts across the participants for each pathway, thereby producing population probability maps, in which voxel values represented the number of participants. The population probability maps with thresholding (at least half of the participants) were smoothed and shown using MRIcroN software[2].

ROI Selection

We basically followed the ROI selection procedures in our previous study (Yamamoto and Sakai, 2016); we determined one-dimensional ROIs (in an antero-posterior direction) at the portion with the most uniform *thickness*, and showed that the appropriate length, i.e., *ROI size*, minimizing the individual variances of thickness was 20 and 15 mm for the Arcuate and IFOF, respectively. We defined the thickness of a pathway as the number of voxels at a coronal section (voxel size, 1 mm^2), because the relatively straight portions of these two pathways (i.e., the portions without branches) were nearly horizontal. We first measured the thickness of each participant's Arcuate for the length of 35 mm ($y = -40 \sim -6$; the candidate region for ROI), excluding the branching or curved portions. Within these candidate regions, we then calculated the standard deviation (SD) of the thickness across a tract segment of 20 mm, and slid the segment. At each position, the *averaged SD* was obtained among all the participants of the three groups. We selected the segment with a minimal averaged SD as an ROI in each hemisphere, thereby extracting the region with the most uniform thickness across participants. Using these procedures, we objectively selected ROIs at the same position in the MNI space, thereby minimizing individual variabilities in the ROI selection.

In regard to the IFOF, we set a candidate region with a length of 70 mm and selected a 15-mm-long ROI in each hemisphere, in accordance with our previous study. We measured the thickness of each participant's IFOF for $y = -75 \sim -6$, excluding the branching or curved portions. Within these candidate regions, we calculated the SD of the thickness across a tract segment of 15 mm, and slid the segment. At each position, the averaged SD was obtained among all the participants. We selected the segment with a minimal averaged SD as an ROI in each hemisphere.

We next examined the correlation with individual accuracy of Syn/Spe to clarify which aspects of L2 abilities were related to structural property that showed any performance-related group differences. For the analyses within a single group, we aimed to examine individual variances in a pathway, thereby employing individually selected ROIs on the normalized tracts. We selected ROIs where the thickness was most uniform for the tract of each participant in the MNI space. As described above, the thickness of the Arcuate was measured for the candidate region with a length of 35 mm ($y = -40 \sim -6$). Within these candidate regions, we calculated the SD of the thickness across a tract segment of

20 mm, and slid the segment. We selected the segment with a minimal SD as an ROI.

In order to avoid the potential uncertainty of FA near the peripheral regions of fiber tracts, we set the thresholding (FA $\geq$ 0.2) in accordance with previous literature. The resultant FA maps were normalized using the affine and non-linear transformation with FLIRT and FNIRT. Within each ROI, we calculated the mean FA for each group, and examined the group differences. Statistical analyses were performed using R software. We used the packages cocor[3] for comparing correlations, ppcor[4] for partial correlations, and pwr[5] for power analyses, as well as lme4[6] and car[7] for linear mixed-effects models (Koerner and Zhang, 2017).

RESULTS

Behavioral Data

Task performances of Syn and Spe in the Junior (High), Senior (Low), and Senior (High) groups are shown in **Figure 1**. Regarding the accuracy (**Figure 1A**), a two-way repeated measures analysis of variance (rANOVA) [group × task (Syn, Spe)] showed significant main effects of group [$F_{(2,44)} = 47$, $P < 0.0001$] and task [$F_{(1,44)} = 169$, $P < 0.0001$] without interaction [$F_{(2,44)} = 0.4$, $P = 0.6$]. A linear mixed-effects model analysis [fixed effects: group, task; random effects: subject] confirmed significant effects of group [$\chi^2(2) = 68$, $P < 0.0001$] and task [$\chi^2(1) = 219$, $P < 0.0001$]. Regarding RTs (**Figure 1B**), an rANOVA showed significant main effects of group [$F_{(2,44)} = 6.3$, $P < 0.005$] and task [$F_{(1,44)} = 116$, $P < 0.0001$] with a significant interaction [$F_{(2,44)} = 4.2$, $P < 0.05$]. A linear mixed-effects model analysis confirmed significant effects of group [$\chi^2(2) = 13$, $P < 0.005$] and task [$\chi^2(1) = 102$, $P < 0.0001$]. One-way ANOVAs did not show a significant effect of gender for the accuracy or RTs of either task [$F_{(1,45)} < 0.2$, $P > 0.7$]. According to the *t*-tests (Bonferroni-corrected for three comparison pairs, significance level at $\alpha = 0.017$), we confirmed that the Senior (High) group had significantly higher accuracy and shorter RTs than the other two groups in both tasks (effect size: $d > 0.90$, statistical power: $1 - \beta > 0.73$), indicating that the Senior (High) group had better syntactic abilities and word knowledge than the other two groups.

If a behavioral parameter reflected factors commonly involved in both Syn and Spe, such as reading proficiency and task difficulty, that parameter would be correlated between Syn and Spe among the participants (see Yamamoto and Sakai, 2016). For instance, RTs of a slow reader would be relatively long among the participants for both tasks. Conversely, if a behavioral parameter was not correlated between the two tasks, that parameter would reflect, at very least, factors distinctly required in each task. Thus we analyzed the correlations among accuracy of the two

---

[2]http://people.cas.sc.edu/rorden/mricron/index.html

[3]https://cran.r-project.org/web/packages/cocor/

[4]https://cran.r-project.org/web/packages/ppcor/

[5]https://cran.r-project.org/web/packages/pwr/

[6]https://cran.r-project.org/web/packages/lme4/

[7]https://cran.r-project.org/web/packages/car/

FIGURE 1 | Behavioral data of Syn and Spe tasks for the three groups. (A) The accuracy of Syn and Spe tasks for the three groups of high school students, Junior (High), Senior (Low), and Senior (High), are indicated by bars with white, gray, and black shades, respectively. Error bars indicate the standard error of the mean (SEM) for the participants, and asterisks denote the significant differences (Bonferroni-corrected $P < 0.05$). **(B)** The reaction times (RTs) of the Syn and Spe tasks for the three groups are indicated by dots corresponding to the white, gray, and black shades in **(A)**. The Senior (High) group showed significantly higher L2 abilities, i.e., higher accuracy and shorter RTs, while the Junior (High) and Senior (Low) groups showed no significant difference in the accuracy or RTs in either task. **(C,D)** Independence of the accuracy of the Syn and Spe tasks. Individual behavioral parameters were plotted to compare the two tasks, as shown by the plotted dots, whose groups correspond to the shades in **(A)**. The dotted line indicates an inclusion criterion of Spe. The accuracy of Syn and Spe showed no significant correlation **(C)**, while the RTs of Syn and Spe were highly correlated **(D)**.

tasks, as well as among RTs. The accuracy of Syn and Spe was not significantly correlated ($r = 0.27$, $P = 0.07$; **Figure 1C**); it showed a week correlation among the three groups, as the Senior (High) group showed the higher accuracy of Syn/Spe than the other two groups, but when examined in each group, such significant positive correlation between the accuracy of Syn and Spe was not observed in any of the three groups ($P > 0.3$). These results suggest that the accuracy of Syn and Spe mainly reflected abilities distinctly required for each task. In contrast, the RTs of Syn and Spe were highly correlated ($r = 0.82$, $P < 0.0001$; **Figure 1D**). Indeed, the RTs of Syn and Spe were positively correlated in all three of the groups ($r > 0.7$, $P < 0.005$). These results indicate that the RTs of each participant were related to general cognitive processes common to both tasks. Moreover, the correlation coefficient between RTs of Syn and Spe was significantly larger compared to that between the accuracy of Syn and Spe ($Z = 4.4$, $P < 0.005$), confirming the differential natures

of these two behavioral measurements. These results suggest that a participant's *accuracy* of each task reflected individual abilities employed for each task. On the other hand, RTs were metrics that reflected more general cognitive abilities, such as reading proficiency. Once a structural property which showed any group difference related to L2 acquisition was found, we further examined its correlation with the accuracy of Syn or Spe to reveal which ability was dominantly reflected. Moreover, we also examined its correlation with the RTs to reveal if this structural property was related to general cognitive abilities.

We also examined the verbal fluency data in L1, and found a consistent trend among groups with the Syn and Spe performances. Behavioral data of the verbal fluency task are shown in **Table 1**. A one-way ANOVA showed a significant effect of group [$F(2,44) = 4.0$, $P < 0.05$]. The Senior (High) group produced significantly larger numbers of words than the Junior (High) group [$t(23) = 2.8$, $P < 0.005$, $d = 1.1$, $1 − \beta = 0.82$],

surviving Bonferroni correction for the three comparison pairs (significance level at $\alpha = 0.017$); the Senior (High) group also produced larger numbers of words than the Senior (Low) group [$t(31) = 2.0$, $P = 0.03$, $d = 0.7$, $1 - \beta = 0.52$]. For the performance-matched groups, i.e., the Junior (High) group and the Senior (Low) group, there was no significant difference [$t(30) = 0.8$, $P = 0.2$]. To examine how syntactic abilities in L2 were related to overall proficiency in L1, we further examined the correlations between the verbal fluency in L1 and the accuracy of Syn for each group. While correlations were not significant for the Senior (High) ($r = 0.23$, $P = 0.4$) and Senior (Low) ($r = 0.10$, $P = 0.7$) groups, a significant positive correlation was present for the Junior (High) group ($r = 0.68$, $P < 0.01$). These results suggest that syntactic abilities in L2 may depend on L1 proficiency in early acquisition stages (DOE ≈ 1.5 years), but this relationship became weaker in later acquisition stages (DOE ≈ 4.5 years).

Group Differences along the Tract of the Arcuate

The Arcuate and IFOF were successfully tracked in both hemispheres for every participant, and the tracking was basically similar among groups. In all groups, the Arcuate connected the frontal and temporal regions with a similar curvature, and the IFOF connected the frontal and occipital/temporal regions through a narrower portion near the external capsule (**Figure 2**). While these overall characteristics were the same for the left and right hemispheres, the left Arcuate was thicker than the right Arcuate. From both lateral and top views, the Arcuate was consistently thicker and extended more anteriorly in the left hemisphere. The Arcuate of the Senior (High) and Senior (Low) groups was thicker than that of the Junior (High) group, which was evident from both views. Such group differences were not observed for the IFOF in either hemisphere.

To examine the overall profile while excluding highly variable regions among individuals, we first set the candidate regions for ROI selection: the region of 35 mm for the Arcuate ($y = -40 \sim -6$), and the region of 70 mm for the IFOF ($y = -75 \sim -6$) (**Figure 2**). As regards the Arcuate, the candidate region was selected where the pathway was relatively straight, excluding the curved or branching portions. The candidate region for the IFOF was also selected where the pathway was straight, excluding the narrower portion near the external capsule, resulting in a longer candidate region than for the Arcuate. The thickness of the Arcuate in the Senior (High) and Senior (Low) groups was significantly larger than that of the Junior (High) group throughout the candidate region in both hemispheres, as indicated by the non-overlapping error bars (mean ± SEM) (**Figure 3A**). In contrast, for the thickness of the IFOF, no clear group difference was observed in either hemisphere. For most of the candidate regions in the Arcuate and IFOF, the thickness was basically uniform in both hemispheres. We selected one-dimensional ROIs, where the thickness was most uniform (see the "Materials and Methods" section). Because the thickness was independent of FA, our ROIs were free from sampling bias of extracting regions with particularly higher or lower FA. We also plotted FA in the same regions in each

hemisphere, and found that FA of the left Arcuate in the Senior (High) group was higher than those of the other two groups, especially in the anterior regions, as indicated by the non-overlapping error bars (**Figure 3B**). In contrast, FA of the right Arcuate did not show such clear group differences. In regard to the IFOF, however, no clear group difference was observed. In these ROIs, FA in the Arcuate showed some modulations, while FA in the IFOF showed relatively sharp antero-posterior changes. Nevertheless, the overall tendency of thickness or FA was similar among the three groups throughout the candidate regions in both hemispheres.

Distinct Group Differences in the Structural Properties of the Left Arcuate

For all the participants, the ROIs were placed at the same position of each pathway in the MNI space. Within these ROIs, we calculated the mean thickness of the Arcuate, and confirmed the leftward laterality (**Figure 2**). A two-way rANOVA [group × hemisphere (left, right)] indicated significant main effects of group [$F(2,44) = 4.9$, $P = 0.01$] and hemisphere [$F(1,44) = 27$, $P < 0.001$], without an interaction [$F(2,44) = 0.3$, $P = 0.7$] (**Figure 4A**). Considering the potential relationships between the thickness in the left and right hemispheres of the same participants, we used a linear mixed-effects model analysis [fixed effects: group, hemisphere; random effects: subject], confirming significant effects of group [$\chi^2(2) = 9.8$, $P = 0.007$] and hemisphere [$\chi^2(1) = 28$, $P < 0.0001$]. Indeed, in all of the three groups, the mean thickness in the ROI of the Arcuate was significantly larger in the left than in the right hemisphere ($P \leq 0.02$) (one-sided t-tests). As regards the group differences, the mean thickness of the left Arcuate for the Senior (High) group was significantly larger than that of the Junior (High) group [$t(27) = 3.1$, $P = 0.002$, $d = 1.2$, $1 - \beta = 0.94$] (one-sided t-tests); the mean thickness for the Senior (Low) group was also larger than that of the Junior (High) group [$t(30) = 1.8$, $P = 0.04$, $d = 0.7$, $1 - \beta = 0.59$]. There was no significant difference between the Senior (High) and Senior (Low) groups [$t(31) = 0.86$, $P = 0.2$]. Regarding the mean thickness of the right Arcuate, the thickness for the Senior (High) group was significantly larger than that for the Junior (High) group [$t(27) = 3.0$, $P = 0.003$, $d = 1.2$, $1 - \beta = 0.93$]. There was no significant difference between the Senior (Low) and Junior (High) groups [$t(30) = 1.3$, $P = 0.10$] or between the Senior (High) and Senior (Low) groups [$t(31) = 1.6$, $P = 0.06$]. No significant effect of gender was observed in either hemisphere, according to one-way ANOVAs [$F(1,45) < 0.4$, $P > 0.5$].

Based on the results of FA shown in **Figure 3B**, we focused on group differences in the left Arcuate. For all voxels (with FA values of 0.2 or higher) within the tract at the ROI of the left Arcuate, we calculated the mean FA, which was clearly higher in the Senior (High) group than in the other two groups (**Figure 4B**). A one-way ANOVA for FA in the left Arcuate indicated a significant main effect of group [$F(2,44) = 3.3$, $P < 0.05$]. More specifically, FA in the left Arcuate for the Senior (High) group was significantly higher than that for the Junior (High) group [$t(27) = 2.2$, $P = 0.016$, $d = 0.9$, $1 - \beta = 0.75$];
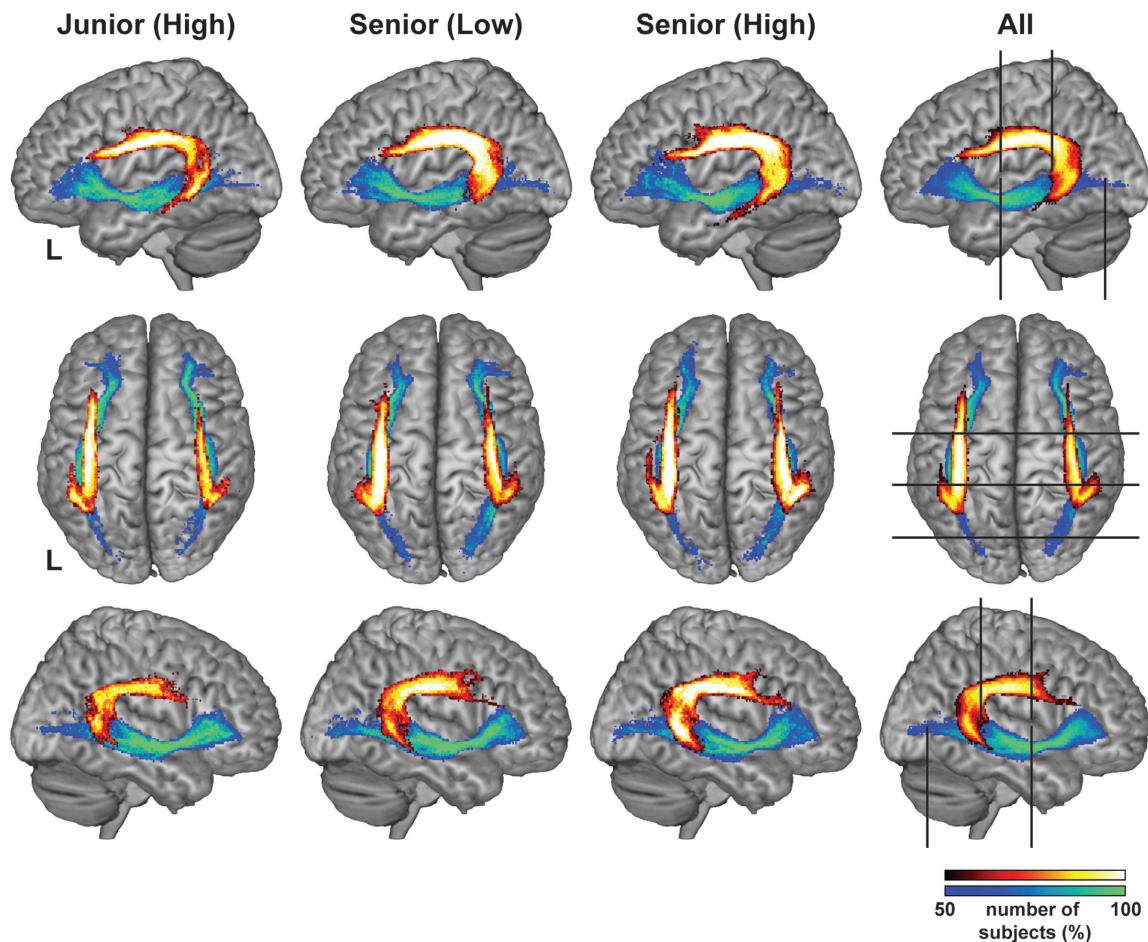
FIGURE 2 | The Arcuate and IFOF reconstructed with diffusion MRI. The probability maps of population for the Arcuate and IFOF in the MNI space. The pathways were thresholded to show the results of the tracking present in more than half of the participants in each group. The color scales denote the number of participants (%): black–red to yellow–white for the Arcuate, and blue to green for the IFOF. Note that the Arcuate in the Senior (High) and Senior (Low) groups was clearly thicker than that in the Junior (High) group, which was evident from both left lateral (top row) and top (middle row) views. Similar tendency was observed in the right hemisphere from the right lateral (bottom row) view. To examine the overall profile while excluding highly variable regions among individuals, we first set candidate regions for selecting regions of interest (ROIs): a region of 35 mm for the Arcuate ($y = -40 \sim -6$), and a region of 70 mm for the IFOF ($y = -75 \sim -6$). The right-most panel shows the candidate regions, which are bounded by the pairs of black lines. We further selected ROIs at the portion with the most uniform thickness in each pathway to quantify the Arcuate and IFOF (see the "Materials and Methods" section), as shown in **Figure 3**.

FA in the left Arcuate for the Senior (High) group was also higher than that for the Senior (Low) group [$t(31) = 2.2$, $P = 0.019$, $d = 0.8$, $1 - \beta = 0.72$]. There was no significant difference between the Senior (Low) and Junior (High) groups [$t(30) = 0.02$, $P = 0.5$]. Regarding FA in the ROI of the right Arcuate, a one-way ANOVA did not show a significant main effect of group [$F(2,44) = 2.0$, $P = 0.15$]. No significant effect of gender was observed for FA in either hemisphere, according to one-way ANOVAs [$F(1,45) < 0.2$, $P > 0.7$].

In regard to the IFOF, neither the mean thickness nor mean FA in the ROI showed a group difference (**Figures 4C,D**). A one-way ANOVA for the thickness did not show a significant main effect of group in either hemisphere [$F(2,44) < 1.7$, $P > 0.8$]. In addition, a one-way ANOVA for FA in the IFOF did not show a significant main effect of group in either hemisphere [$F(2,44) < 2.6$, $P > 0.05$]. Major structural properties, i.e., the

thickness and FA, in the ROI of IFOF in both hemispheres were similar among the three groups.

FA in the Left Arcuate Was Selectively Correlated with the Accuracy of Syn

For FA in the left Arcuate, which showed group differences between the Senior (High) and the other two groups (see **Figure 4B**), we examined what aspect of L2 abilities was actually related to FA. We performed the following analyses for the Senior (High) group, whose task performances were higher and thus most reliable for dissociating the linguistic abilities required by Syn or Spe. ROIs were selected for each participant as described in the "Materials and Methods" section. We performed partial correlation analyses between the standardized accuracy of Syn and the standardized FA in the left Arcuate, removing the effects
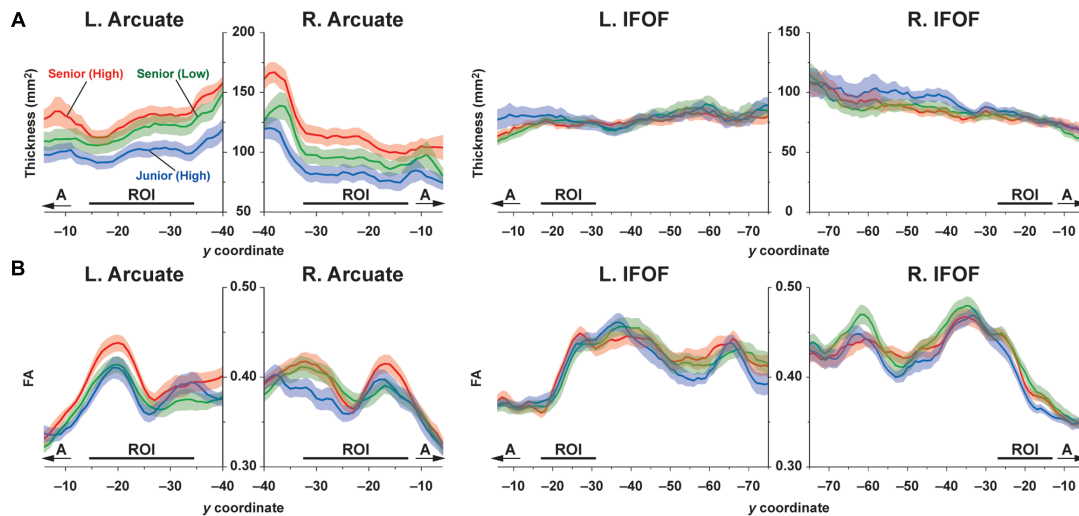
FIGURE 3 | Profiles of the thickness/FA of the dorsal and ventral pathways in each group. (A) The profiles of thickness for the Arcuate and IFOF. **(B)** The profiles of FA for the Arcuate and IFOF. The thickness and FA were averaged among the participants in each of the Junior (High), Senior (Low), and Senior (High) groups, as shown in blue, green, and red, respectively. Data are shown for the range of $y = -40 \sim -6$ for the Arcuate and $y = -75 \sim -6$ for the IFOF (bounded by the pairs of black lines in **Figure 2**) in the MNI space. Note that the thickness of the Arcuate in the Senior (High) group is larger than that in the Junior (High) group in both hemispheres. Moreover, FA in the *left* Arcuate in the Senior (High) group is higher than those in the other two groups. No clear group difference was found for the IFOF in either hemisphere. The SEMs in each group are shown as shaded bands in each color. The positions of ROIs are indicated by the black lines above the axes. A, anterior.
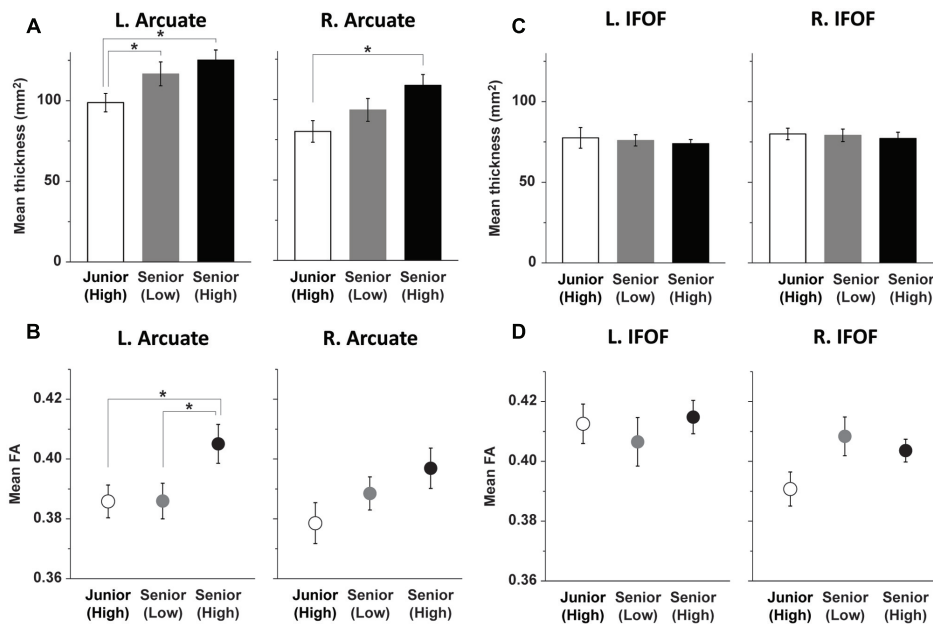
FIGURE 4 | Group differences in the structural properties of the left Arcuate. (A) Age-related group differences on the thickness of the left dorsal pathway. **(B)** Performance-related group differences on FA in the left dorsal pathway. **(C,D)** The ventral pathway without group differences. The Junior (High), Senior (Low), and Senior (High) groups are indicated by bars/dots with white, gray, and black shades, respectively. Error bars indicate the SEM for the participants, and asterisks denote statistical differences at $P < 0.05$ (*post hoc t*-test after ANOVA).

of the accuracy of Spe, LQ, and gender. Regarding the accuracy of Syn, we found a significant correlation with FA in the left Arcuate ($r = 0.61$, $P = 0.03$) (**Figure 5A**). In contrast, the accuracy of Spe was not significantly correlated with FA in the left Arcuate

($r = 0.40$, $P = 0.2$), according to the partial correlation analysis removing the effects of the accuracy of Syn, LQ, and gender (**Figure 5B**). In addition, no significant correlation was found between FA in the left Arcuate and verbal fluency in L1 ($r = 0.24$,

FIGURE 5 | FA in the left Arcuate was selectively correlated with the accuracy of Syn. (A) The scatter plots of the standardized accuracy of Syn and the standardized mean FA in the left Arcuate. The ROIs in the left Arcuate were determined for each participant in the Senior (High) group. The effects of the standardized accuracy of Spe, gender, and laterality quotient (LQ) were removed. **(B)** The scatter plots of the standardized accuracy of Spe and the standardized mean FA in the left Arcuate. The effects of the standardized accuracy of Syn, gender, and LQ were removed. FA in the left Arcuate was correlated with the accuracy of Syn, but not with the accuracy of Spe.

$P = 0.4$) in the partial correlation analysis removing the effects of LQ and gender. Moreover, no significant correlation was found between FA in the left Arcuate and the RTs of Syn ($r = 0.17$,

$P = 0.6$), according to a partial correlation analysis removing the effects of the RTs of Spe, LQ, and gender. These results indicate that increased FA in the left Arcuate for the Senior (High) group was related mainly to the enhanced syntactic abilities in L2, irrespective of L1 performances or other general measures examined here.

DISCUSSION

In the three groups of high school students, we obtained the following results. First, performance-related group differences were found only for FA in the left Arcuate. More specifically, the Senior (High) group, who had higher L2 abilities, showed higher FA in the left Arcuate than the Senior (Low) and Junior (High) groups (**Figures 3B**, **4B**). Moreover, the mean FA in the ROI of the left Arcuate of the Senior (High) group was significantly correlated with the accuracy of Syn (**Figure 5A**), but not with the accuracy of Spe (**Figure 5B**) or with verbal fluency in L1, indicating that increased FA in the left Arcuate was related mainly to the enhanced syntactic abilities in L2. Secondly, age-related group differences were found for the thickness of the left Arcuate. The thickness for the Senior (High) and Senior (Low) groups was larger than that for the Junior (High) group (**Figures 3A**, **4A**), indicating that the left Arcuate was still developing in adolescence. Thirdly, these differential performance-related and age-related signatures were evident on the left Arcuate alone, in contrast to the right Arcuate that showed only mild differences in thickness (**Figures 3A**, **4A**), and to the bilateral IFOF that lacked either signature (**Figures 3**, **4C,D**). To the best of our knowledge, our study is the first to report that plasticity during the adolescent years was markedly different between the dorsal and ventral language-related pathways, the former of which was related to syntactic abilities. In summary, we showed that the left dorsal pathway, which has been reported to be more immature than the ventral pathway from early infancy, continued to grow thicker with increasing age at least until the adolescent years. Further, by showing the group difference between the Senior (High) and Senior (Low) groups, whose DOEs were the same, we revealed that performance differences were reflected in the FA in the left dorsal pathway, independent of age and the duration of experience. These results indicate that the left dorsal pathway is a major neural network supporting syntactic abilities.

Within the same part of the left dorsal pathway, we observed the dissociated performance-related and age-related group differences on FA and the thickness, respectively. These results provide new insights into the plasticity of these two distinct structural properties in the left Arcuate. FA and the volume of the Arcuate have been reported to change during the adolescent years (Lebel and Beaulieu, 2011), but it has remained unclear whether or not these properties change in accordance with the development of specific abilities. Our present results suggest that not age or maturation *per se* but the gradual improvement of performances in syntactic abilities during adolescence was related to FA. The increased FA in accordance with certain learning or L2 exposure has been reported also in other pathways. For instance,

an increase in FA in the genu of the corpus callosum was correlated with the overall L2 performances after a 9-month intensive course of spoken and written Chinese (Schlegel et al., 2012). Moreover, an increase in FA in the white matter underlying the right IFG was correlated with vocabulary test scores after a 16-week period of vocabulary training in English (Hosoda et al., 2013). In our previous study, focusing on students with better orthographic knowledge (80% or higher accuracy in the Spe task), we showed that individual differences in the accuracy of Syn were reflected in FA in the left Arcuate (Yamamoto and Sakai, 2016). In the present study, analyzing students with a wider range of L2 abilities and ages, we demonstrated that FA in the left Arcuate was higher in students with higher L2 abilities than in students with lower abilities at the same age, whose FA was not significantly different from performance-matched younger students. Further, we showed that FA in the left Arcuate of the Senior (High) group was positively correlated with the accuracy of the syntactic task, but not with the accuracy of the spelling task. These results indicate that the higher FA in the left dorsal pathway was mainly related to the higher syntactic abilities in L2. As described in the "Materials and Methods" section, the Syn task was designed to test argument structures and related syntactic knowledge in English, and cannot be correctly answered by semantical/pragmatic cues, or by internally translating English sentences into Japanese. Indeed, the accuracy of Syn and L1 verbal fluency was not significantly correlated in the Senior (High) and Senior (Low) groups. Future studies with tasks that can accurately assess individual syntactic abilities in L1 will be needed to reveal how L1 syntactic abilities, as well as neural networks supporting these abilities, facilitate (or hamper) the acquisition of a new language. Moreover, it would be also important to elucidate how the development in other aspects of linguistic abilities is related to the plasticity in the left dorsal pathway.

Another important issue is whether non-cognitive abilities are related to the development of the language-related structures. A previous study on L2 learning in infants indicated the impact of social interaction, by showing that infants who received "live-person" sessions on Mandarin performed significantly better on Mandarin phonetic perception tests, compared to infants who received the identical information via television or audiotape but showed no learning effects (Kuhl et al., 2003). These social factors, as well as other non-cognitive factors such as motivation and self-confidence, may have impacts on L2 acquisition in adolescent students and adults as well (Dörnyei, 2003; Gardner, 2010). One of the directions of future studies would be to investigate how such individual traits affect one's acquisition experience, which may further modify functional and structural networks. It would also be important to dissociate the effects of non-cognitive abilities on language-related networks from those of cognitive abilities on the left dorsal pathway, whose critical involvement in linguistic information processing has been shown by previous studies (Ohta et al., 2013; Yamamoto and Sakai, 2016), and confirmed in the present study.

Here, we showed that the thickness of the left Arcuate was larger in the Senior (High) and Senior (Low) groups than in the Junior (High) group, indicating age-related group differences. This macro-structural property of the left Arcuate had plasticity

associated with age, which may reflect biological maturation, as well as common and specific experiences that students underwent during the ages of 13–17. A recent longitudinal study reported that children's structural connectivity, which was obtained before they learned to read (i.e., before age 5), predicted the spatial profile of functional activations in the Visual Word Form Area (VWFA) after they learned to read (i.e., after age 8), suggesting that connectivity precedes the functional development (Saygin et al., 2016). These results raise an interesting hypothesis that experience interacts with the micro-structural development of a pathway whose structural connectivity has already developed. A more detailed picture of the developmental mechanism of language-related structures would be provided by closely linking the development of linguistic abilities and that of white matter pathways, as well as the structures of connected cortical regions, which can be examined by such methods as VBM and the myelin mapping technique (Glasser and Van Essen, 2011). Our present study showed both age-related and performance-related group differences in adolescent participants, suggesting the importance of future longitudinal studies with various structural and functional measurements.

In addition to the fronto-temporal segment examined here, the arcuate fasciculus is also composed of the fronto-parietal and temporo-parietal segments (Catani et al., 2005). Thiebaut de Schotten et al. (2014) have suggested that the temporo-parietal segment has plasticity associated with learning even in adults, based on their finding that its FA was higher in a group of ex-illiterates, who lacked access to schools during childhood for social reasons and learned to read during adulthood, in contrast to a group of illiterates, who never learned to read. They also showed that FA was correlated with activations in the two regions, i.e., the left VWFA and planum temporale, connected to the angular/supramarginal gyri by this segment. As the authors of this study discussed in their subsequent paper (Dehaene et al., 2015), these results might have been influenced by multiple factors, including the motivations, self-confidence, socioeconomic status, and professions of the participants. Indeed, differences in FA between ex-illiterates and illiterates might be present before ex-illiterates voluntarily begin to learn reading. Future studies should attempt to verify whether or not "learning to read improves the structure" or not. To examine such causal changes in the brain, longitudinal studies that track neural indices and behavioral measures, together with multiple regression analyses that de-correlate confounding variables, would be critical (Galván, 2010; Dehaene et al., 2015). Note that our present study does not intend to suggest causal influences of L2 acquisition on structural properties. Rather, we demonstrated here that the performance-related group differences, which were separated from the duration of L2 experience (i.e., DOE) as well as from the age-related group differences, were predicted by FA in the left dorsal pathway.

AUTHOR CONTRIBUTIONS

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Brauer, J., Anwander, A., Perani, D., and Friederici, A. D. (2013). Dorsal and ventral pathways in language development. *Brain Lang.* 127, 289–295. doi: 10.1016/j.bandl.2013.03.001

Brown, T. T., Lugar, H. M., Coalson, R. S., Miezin, F. M., Petersen, S. E., and Schlaggar, B. L. (2005). Developmental changes in human cerebral functional organization for word generation. *Cereb. Cortex* 15, 275–290. doi: 10.1093/cercor/bhh129

Catani, M., Jones, D. K., and Ffytche, D. H. (2005). Perisylvian language networks of the human brain. *Ann. Neurol.* 57, 8–16. doi: 10.1002/ana.20319

Dehaene, S., Cohen, L., Morais, J., and Kolinsky, R. (2015). Illiterate to literate: behavioural and cerebral changes induced by reading acquisition. *Nat. Rev. Neurosci.* 16, 234–244. doi: 10.1038/nrn3924

Dörnyei, Z. (2003). Attitudes, orientations, and motivations in language learning: advances in theory, research, and applications. *Lang. Learn.* 53, 3–32. doi: 10.1111/1467-9922.53222

Friederici, A. D. (2011). The brain basis of language processing: from structure to function. *Physiol. Rev.* 91, 1357–1392. doi: 10.1152/physrev.00006.2011

Fuhrmann, D., Knoll, L. J., and Blakemore, S.-J. (2015). Adolescence as a sensitive period of brain development. *Trends Cogn. Sci.* 19, 558–566. doi: 10.1016/j.tics.2015.07.008

Galván, A. (2010). Neural plasticity of development and learning. *Hum. Brain Mapp.* 31, 879–890. doi: 10.1002/hbm.21029

Gardner, R. C. (2010). *Motivation and Second Language Acquisition: The Socio-Educational Model.* New York, NY: Peter Lang.

Giorgio, A., Watkins, K. E., Chadwick, M., James, S., Winmill, L., Douaud, G., et al. (2010). Longitudinal changes in grey and white matter during adolescence. *Neuroimage* 49, 94–103. doi: 10.1016/j.neuroimage.2009.08.003

Glasser, M. F., and Van Essen, D. C. (2011). Mapping human cortical areas in vivo based on myelin content as revealed by T1- and T2-weighted MRI. *J. Neurosci.* 31, 11597–11616. doi: 10.1523/jneurosci.2180-11.2011

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., et al. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proc. Natl. Acad. Sci. U.S.A.* 101, 8174–8179. doi: 10.1073/pnas.0402680101

Hickok, G., and Poeppel, D. (2007). Opinion - The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402. doi: 10.1038/nrn2113

Hosoda, C., Tanaka, K., Nariai, T., Honda, M., and Hanakawa, T. (2013). Dynamic neural network reorganization associated with second language vocabulary acquisition: a multimodal imaging study. *J. Neurosci.* 33, 13663–13672. doi: 10.1523/jneurosci.0410-13.2013

Koerner, T., and Zhang, Y. (2017). Application of linear mixed-effects models in human neuroscience research: a comparison with Pearson correlation in two auditory electrophysiology studies. *Brain Sci.* 7, 1–11. doi: 10.3390/brainsci7030026

Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nat. Rev. Neurosci.* 5, 831–843. doi: 10.1038/nrn1533

Kuhl, P. K., Tsao, F.-M., and Liu, H.-M. (2003). Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. *Proc. Natl. Acad. Sci. U.S.A.* 100, 9096–9101. doi: 10.1073/pnas.1532872100

Lebel, C., and Beaulieu, C. (2011). Longitudinal development of human brain wiring continues from childhood into adulthood. *J. Neurosci.* 31, 10937–10947. doi: 10.1523/jneurosci.5302-10.2011

Lenroot, R. K., Gogtay, N., Greenstein, D. K., Wells, E. M., Wallace, G. L., Clasen, L. S., et al. (2007). Sexual dimorphism of brain developmental trajectories during childhood and adolescence. *Neuroimage* 36, 1065–1073. doi: 10.1016/j.neuroimage.2007.03.053

Li, P., Legault, J., and Litcofsky, K. A. (2014). Neuroplasticity as a function of second language learning: anatomical changes in the human brain. *Cortex* 58, 301–324. doi: 10.1016/j.cortex.2014.05.001

Mohades, S. G., Van Schuerbeek, P., Rosseel, Y., Van De Craen, P., Luypaert, R., and Baeken, C. (2015). White-matter development is different in bilingual and monolingual children: a longitudinal DTI study. *PLoS ONE* 10:e0117968. doi: 10.1371/journal.pone.0117968

Nauchi, A., and Sakai, K. L. (2009). Greater leftward lateralization of the inferior frontal gyrus in second language learners with higher syntactic abilities. *Hum. Brain Mapp.* 30, 3625–3635. doi: 10.1002/hbm.20790

Ohta, S., Fukui, N., and Sakai, K. L. (2013). Syntactic computation in the human brain: the degree of merger as a key factor. *PLoS ONE* 8:e56230. doi: 10.1371/journal.pone.0056230

Oldfield, R. C. (1971). The assessment and analysis of handedness: the edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4

Perani, D., Saccuman, M. C., Scifo, P., Anwander, A., Spada, D., Baldoli, C., et al. (2011). Neural language networks at birth. *Proc. Natl. Acad. Sci. U.S.A.* 108, 16056–16061. doi: 10.1073/pnas.1102991108

Sakai, K. L., Nauchi, A., Tatsuno, Y., Hirano, K., Muraishi, M., Kimura, M., et al. (2009). Distinct roles of left inferior frontal regions that explain individual differences in second language acquisition. *Hum. Brain Mapp.* 30, 2440–2452. doi: 10.1002/hbm.20681

Saygin, Z. M., Osher, D. E., Norton, E. S., Youssoufian, D. A., Beach, S. D., Feather, J., et al. (2016). Connectivity precedes function in the development of the visual word form area. *Nat. Neurosci.* 19, 1250–1255. doi: 10.1038/nn.4354

Schlaggar, B. L., Brown, T. T., Lugar, H. M., Visscher, K. M., Miezin, F. M., and Petersen, S. E. (2002). Functional neuroanatomical differences between adults and school-age children in the processing of single words. *Science* 296, 1476–1479.

Schlegel, A. A., Rudelson, J. J., and Tse, P. U. (2012). White matter structure changes as adults learn a second language. *J. Cogn. Neurosci.* 24, 1664–1670. doi: 10.1162/jocn_a_00240

Schwarz, C. G., Reid, R. I., Gunter, J. L., Senjem, M. L., Przybelski, S. A., Zuk, S. M., et al. (2014). Improved DTI registration allows voxel-based analysis that outperforms tract-based spatial statistics. *Neuroimage* 94, 65–78. doi: 10.1016/j.neuroimage.2014.03.026

Thiebaut de Schotten, M., Cohen, L., Amemiya, E., Braga, L. W., and Dehaene, S. (2014). Learning to read improves the structure of the arcuate fasciculus. *Cereb. Cortex* 24, 989–995. doi: 10.1093/cercor/bhs383

van Gelderen, A., Schoonen, R., Stoel, R. D., De Glopper, K., and Hulstijn, J. (2007). Development of adolescent reading comprehension in Language 1 and Language 2: a longitudinal analysis of constituent components. *J. Educ. Psychol.* 99, 477–491. doi: 10.1037/0022-0663.99.3.477

Westlye, L. T., Walhovd, K. B., Dale, A. M., Bjørnerud, A., Due-Tønnessen, P., Engvig, A., et al. (2010). Life-span changes of the human brain white matter: diffusion tensor imaging (DTI) and volumetry. *Cereb. Cortex* 20, 2055–2068. doi: 10.1093/cercor/bhp280

Yamamoto, K., and Sakai, K. L. (2016). The dorsal rather than ventral pathway better reflects individual syntactic abilities in second language. *Front. Hum. Neurosci.* 10:295. doi: 10.3389/fnhum.2016.00295

Zöllei, L., Stevens, A., Huber, K., Kakunoori, S., and Fischl, B. (2010). Improved tractography alignment using combined volumetric and surface registration. *Neuroimage* 51, 206–213. doi: 10.1016/j.neuroimage.2010.01.101

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.

"When Music Speaks": Auditory Cortex Morphology as a Neuroanatomical Marker of Language Aptitude and Musicality

Sabrina Turker[1]*, Susanne M. Reiterer[2], Annemarie Seither-Preisler[1,3] and Peter Schneider[4,5]

[1] Centre for Systematic Musicology, University of Graz, Graz, Austria, [2] Department of Linguistics, University of Vienna, Vienna, Austria, [3] BioTechMed-Graz, Graz, Austria, [4] Section of Biomagnetism, Department of Neurology, University Hospital Heidelberg, Heidelberg, Germany, [5] Division of Neuroradiology, University Hospital Heidelberg, Heidelberg, Germany

Recent research has shown that the morphology of certain brain regions may indeed correlate with a number of cognitive skills such as musicality or language ability. The main aim of the present study was to explore the extent to which foreign language aptitude, in particular phonetic coding ability, is influenced by the morphology of Heschl's gyrus (HG; auditory cortex), working memory capacity, and musical ability. In this study, the auditory cortices of German-speaking individuals ($N = 30$; 13 males/17 females; aged 20–40 years) with high and low scores in a number of language aptitude tests were compared. The subjects' language aptitude was measured by three different tests, namely a Hindi speech imitation task (phonetic coding ability), an English pronunciation assessment, and the Modern Language Aptitude Test (MLAT). Furthermore, working memory capacity and musical ability were assessed to reveal their relationship with foreign language aptitude. On the behavioral level, significant correlations were found between phonetic coding ability, English pronunciation skills, musical experience, and language aptitude as measured by the MLAT. Parts of all three tests measuring language aptitude correlated positively and significantly with each other, supporting their validity for measuring components of language aptitude. Remarkably, the number of instruments played by subjects showed significant correlations with all language aptitude measures and musicality, whereas, the number of foreign languages did not show any correlations. With regard to the neuroanatomy of auditory cortex, adults with very high scores in the Hindi testing and the musicality test (AMMA) demonstrated a clear predominance of complete posterior HG duplications in the right hemisphere. This may reignite the discussion of the importance of the right hemisphere for language processing, especially when linked or common resources are involved, such as the inter-dependency between phonetic and musical aptitude.

Keywords: neuroanatomical correlates, language aptitude, musicality, working memory, auditory cortex morphology, Heschl's gyrus

INTRODUCTION

There has always been a fascination with the simple fact that some individuals are strikingly better at doing something, e.g., playing an instrument, singing, or learning a foreign language. It is said that these individuals possess a certain aptitude, i.e., a potential for developing exceptional ability (Gagné, 1995, 2005; Faulkner, 2003; Nardo and Reiterer, 2009; Al-Shabatat, 2013; Stern and Neubauer, 2013). According to Gagné (1995, 2005), aptitude designates the innate property that develops into a certain skill, which is then termed talent (Stern and Neubauer, 2013; Deiglmayr et al., 2017). Individuals with a high aptitude for something mostly put little effort into acquiring a certain skill and need far less time and practice to reach a high achievement or proficiency level in comparison to age-matched peers (Carroll, 1958, 1962, 1990). The concept of language aptitude has gained considerable momentum in the past decades and research has shown that various factors contribute to the overall achievement and proficiency of skills, e.g., environmental influences, personality traits, motivation and other abilities such as musicality or working memory (Ganschow and Sparks, 1995; Dörnyei, 1998, 2006; Sparks and Ganschow, 2001; Brown, 2006; Biedroń and Szczepaniak, 2009; Rota and Reiterer, 2009; Biedroń, 2011a,b, 2012; Sparks et al., 2011; Wen, 2012, 2016; Christiner and Reiterer, 2013; Granena and Long, 2013; Dörnyei and Ryan, 2015; Li, 2015, 2016; Biedroń and Pawlak, 2016a,b; Singleton, 2017; Wen et al., 2017).

Language aptitude is a vague concept challenging to grasp and even more difficult to measure accurately (Li, 2015). In the 1st years after the birth of language aptitude research, it was regarded as an exceptional ability that facilitates foreign language learning in terms that individuals learn a language very quickly and with little effort (Carroll, 1958, 1962, 1973, 1990; Stansfield and Reed, 2004). For a long time, language aptitude was thus defined by the rate of acquisition at which an unknown language was learned. More recent definitions (Robinson, 2005) describe it as the strength of an individual with respect to cognitive abilities especially drawn upon during the learning of foreign languages. In the past years, the focus of foreign language aptitude research has shifted more toward formerly neglected issues, such as the influence and importance of inter-individual differences (Skehan, 1986, 2002; Spolsky, 1995; Dörnyei, 1998, 2006; Robinson, 2001, 2002, 2012; Dörnyei and Skehan, 2003; Biedroń, 2015; Dörnyei and Ryan, 2015; Wen et al., 2017). The four major components of language aptitude claimed by Carroll (1958, 1962), namely (1) Phonetic Coding Ability, (2) Grammatical Sensitivity, (3) Inductive Language Learning Ability, and (4) Rote Learning Ability, are still upheld nowadays. Still, some theoretical advancements have been made and it has been agreed that inductive language learning ability and grammatical sensitivity are most suitably summarized in one category termed language analytic ability (Robinson, 2001, 2002, 2012; Abrahamsson and Hyltenstam, 2008; Kocić, 2010; Biedroń, 2015; Biedroń and Pawlak, 2016a,b; Wen et al., 2017). Moreover, researchers have questioned whether the distinctive components of language aptitude might be more or less influential at different stages and in different contexts of learning (Abrahamsson and Hyltenstam, 2008; Artieda and Muñoz, 2016).

The core factors investigated in this study are working memory, musicality, language aptitude and auditory cortex morphology. Language and music are two inextricably linked concepts that extensively influence one another (Besson and Schön, 2003; Koelsch, 2005; Kraus and Chandrasekaran, 2010; Patel, 2011, 2012; Jäncke, 2012; Chobert and Besson, 2013; Lee and Lin, 2015). A positive correlation between musicality and foreign language aptitude was found in numerous studies, particularly regarding second language pronunciation skills (Schön et al., 2004; Besson et al., 2007; Dogil and Reiterer, 2009; Ludke, 2010; Fonseca-Mora et al., 2011; Christiner and Reiterer, 2013). Whereas Milovanov et al. (2004, 2008, 2009, 2010), Milovanov (2009), Milovanov and Tervaniemi (2011) focused on the successful relationship between musicality and foreign language learning in Finnish native speakers, Vangehuchten et al. (2015) found a significant relationship between English pronunciation skills and musical skills in Spanish native speakers. Dolman and Spring (2014) revealed that excellent skills in specific musical abilities of Japanese learners, such as the discrimination of pitch, loudness, and rhythm, correlate with better pronunciation in the second language (English). Likewise, Slevc and Miyake (2006) found a consistent relationship between musical aptitude and phonological aspects of linguistic ability, but not between syntactic or semantic skills. Similar results were also found for Iranian native speakers (Shabani and Torkeh, 2014). Apart from general musical abilities, singing has also been shown to correlate with foreign language aptitude (Ludke, 2010; Ludke et al., 2014), especially pronunciation aptitude and speech imitation skills (Christiner and Reiterer, 2013, 2015).

Beside the growing interest in musicality as an essential factor for successful foreign language acquisition, the mutual interdependence between working memory (for details, refer to Baddeley and Hitch, 1974, 2000; Baddeley, 2003a,b) and language aptitude has been the focus of most recent research. Due to the strong correlation between the two, some researchers have even gone as far as to claim that working memory capacity is equivalent to language aptitude (Miyake and Friedman, 1998; Sawyer and Ranta, 2001; Wen and Skehan, 2011; Wen, 2016; Wen et al., 2017). Studies on language ability and working memory have confirmed the impact of the latter on numerous linguistic abilities, such as faster and more successful first and second language learning (Ellis and Sinclair, 1996; Miyake and Friedman, 1998; Kormos and Sáfár, 2008; Sáfár and Kormos, 2008; Linck et al., 2013). In other words, those learners who have significantly better working memory skills seem to learn more foreign languages and tend to be more successful (Van den Noort et al., 2006; Biedroń, 2012). However, major issues therein are the differences between specific working memory components, how they can be tested and in how far they relate to the known components of foreign language aptitude (Baddeley, 2003a,b, 2017; Jacquemot and Scott, 2006). Additionally, other studies have questioned the large impact working memory is said to have on language aptitude (Winke, 2013).

While musicality and working memory are mostly treated as clear predictors of foreign language learning ability, the

relationship between brain morphology and language aptitude is far from obvious. The processing of language in the human brain has been a subject of investigation in countless studies (for overviews see Friederici, 2009; Price, 2010, 2012; Xiang, 2012), but very few have actually focused on language aptitude or talent (e.g., Golestani et al., 2007, 2011; Reiterer et al., 2011; Xiang et al., 2012; Hu et al., 2013; Kepinska et al., 2017a,b). In this study, we focus on the neuroanatomy of the auditory cortex given its importance for processing speech. The core region containing the primary auditory cortex is Heschl's gyrus (henceforth always HG), embedded as a transverse gyrus oriented from the insular toward the anterolateral part of the superior temporal lobe. Most humans possess a single or paired HG, the latter in the shape of a common stem or complete posterior duplication (CPD) of HG (Rademacher et al., 1992, 2001; Morosan et al., 2001; Purves et al., 2001; Bear et al., 2006; Hackett, 2009), but HG shows considerable morphological variation between individuals (Heschl, 1878; Galaburda et al., 1978; Rademacher et al., 1992, 2001; Marie et al., 2016), especially in the right hemisphere (Penhune et al., 1996; Schneider et al., 2002, 2005; Seither-Preisler et al., 2014; Serrallach et al., 2016; Benner et al., 2017). Benner et al. (2017) recently found that 90% of musicians had multiplications of HG, mostly on the right side. The right hemisphere has often been claimed to be particularly important for the processing of musical sounds (Zatorre et al., 2002) and less important for speech. Other studies have also suggested that the shape and number of Heschl's gyri may be an indicator for musical skills and auditory-related developmental disorders such as dyslexia (Schneider et al., 2002, 2005; Warrier et al., 2012; Seither-Preisler et al., 2014; Serrallach et al., 2016; Benner et al., 2017). Seither-Preisler et al. (2014) discovered that a large right HG is associated with high musical aptitude in children. Their longitudinal observations revealed that the gross morphology and gray matter volume of different parts of auditory cortex showed a high inter-individual variability, but remained almost perfectly stable throughout the study, lasting for several years. A regression model showed that this neuroanatomical trait was much stronger associated with measures of musical aptitude than with training-related musical expertise (i.e., the amount of previous training). The authors therefore concluded that an enlarged right HG reflects a high predisposition for music which enhances a child's intrinsic motivation to learn an instrument. As a consequence, this leads to high musical expertise and boosts learning-induced neural plasticity. It therefore appears worthwhile to explore possible neuroanatomic markers of language aptitude as well.

With regard to language ability, few studies have addressed the significance of specific language-involved regions, such as HG, for language learning. Kepinska (2017) performed a highly appealing study investigating the neural basis of language analytic ability in high and moderate learners. They found that the more skilled learners drew more from neural resources in the right hemisphere (e.g., right angular gyrus, supramarginal gyrus, superior frontal and middle gyrus, and posterior cingulate), in contrast to the less skilled learners. Golestani et al. (2007) found correlations between an abnormal asymmetry of the planum temporale and poor verbal skills. Golestani et al. (2011) reported that the size of the left inferior frontal gyrus (pars opercularis)

correlated with the years of experience expert phoneticians had. Additionally, they found that the expert phoneticians more frequently had multiple or split HGs in their core auditory cortex in the left hemisphere. In studies by Wong et al. (2007, 2008), English-speaking adults, who were less successful in learning to incorporate foreign pitch patterns in word identification, exhibited smaller volume of HG in the left hemisphere only.

A widespread belief holds that the left temporal lobe is more tuned for the processing of rapid sound stimuli, which consequently leads to a left-hemispheric dominance for speech processing. This hypothesis assigns information processing in short time windows (e.g., phonemes) to the left and longer time windows (e.g., syllables to intonation profiles) to the right hemisphere (Zatorre et al., 2002; Poeppel, 2003; McGettigan and Scott, 2012). Warrier et al. (2012) also support this by reporting that left HG is of greater importance for varying rates of stimulus change, and right HG for music-relevant functions, such as increasing spectral information. However, McGettigan and Scott (2012) question, whether sensitivity for rapid information is sufficient for efficient speech processing for various reasons, e.g., the identification of differences in the duration of consonants and the encoding of supra-segmental information in speech. Most importantly, they argue that the main issue with this hypothesis and widespread belief is 'the assumption that access to phoneme representations is the cardinal aspect of speech perception.'

To summarize, the neuroanatomy of HG has been addressed in various studies focusing on sound and speech perception, but only in few studies dealing with language aptitude. We therefore aim at bridging this gap by exploring the importance of the number of HGs in individuals with high and low language learning abilities.

MATERIALS AND METHODS

Subjects

All participants ($N = 30$; 13 male/17 female) were monolingual German native speakers between 20 and 40 years of age ($M = 26.77$, $SD = 4.95$) and had begun acquiring their second language, English, at 10 ± 1 years of age. All participants were right-handed German bachelor/master students or had achieved positions at an institution of higher education. None of the participants showed any medical condition or neurological disorder. The subjects were paid for participation and provided written informed consent before participating in the experiment. The data were analyzed anonymously.

Language Aptitude Testing

All individuals were classified as high or low aptitude individuals according to two scores, namely an English pronunciation score and a Hindi speech imitation score (both ranging from 0 to 10). The English pronunciation score was based on reading performance of 'The North Wind and the Sun' rated by native speakers. In the Hindi task, participants had to repeat four words and four sentences in Hindi, an unknown language to them. The Hindi imitation and the English pronunciation performances were categorized in a similar fashion – both were rated by

native speakers according to the correctness and quality of the pronunciation/imitation based on a global intuitive impression rating procedure (the rating procedure is detailed in Reiterer et al., 2011 and Jilka, 2009). The inter-rater reliability was 0.96, i.e., very high, because of the unusually high amount of raters ($N = 30$). Thus, the corresponding scores can be considered as highly reliable. The speakers were recorded on a professional speech recording equipment in the sound-proof basement room of the former phonetics laboratory of the Institute of Natural Language Processing, University of Stuttgart. The native speakers were provided with the speech material online and gave ratings from 0 to 10 on an intuitive Likert-scale-like bar for the quality and 'nativelikeness' of the speech material. For the Hindi rating, sound files of Hindi native speakers were added to the rating sample of German speakers (without knowledge of the raters). We used this as additional measure to verify the validity of the rating procedure.

Beside the speech imitation skills (referred to as a measure for phonetic coding ability) and the English pronunciation skills, the Modern Language Aptitude Test (MLAT; Carroll and Sapon, 1957; Ganschow and Sparks, 1995; Dogil and Reiterer, 2009) was used for assessing different components of language aptitude. The parts of the MLAT used were III, IV, and V and the overall total raw score (a combination of the three sub-scores) was further calculated. The three sub-tests provide measures of phonetic coding ability (the ability to differentiate between speech sounds), associative memory (the ability to keep linguistic input in memory and to access this information) and grammatical sensitivity (the ability to understand grammatical relationships and the functions of words in a given context) (Carroll and Sapon, 1957; Carroll, 1958, 1973, 1990); for further details see **Table 1**.

Musicality Assessment

To assess aptitude in the musical domain, the AMMA test (Gordon, 1980, 2001), a well-established tool for administering musical aptitude, was used. It consists of two parts and has been shown to successfully measure pitch and rhythm perception. The subjects were asked to complete both tasks, (1) a rhythm discrimination task and (2) a pitch discrimination task. Furthermore, a questionnaire was used to specify the number

and type of instruments the subjects had learnt in the course of their life.
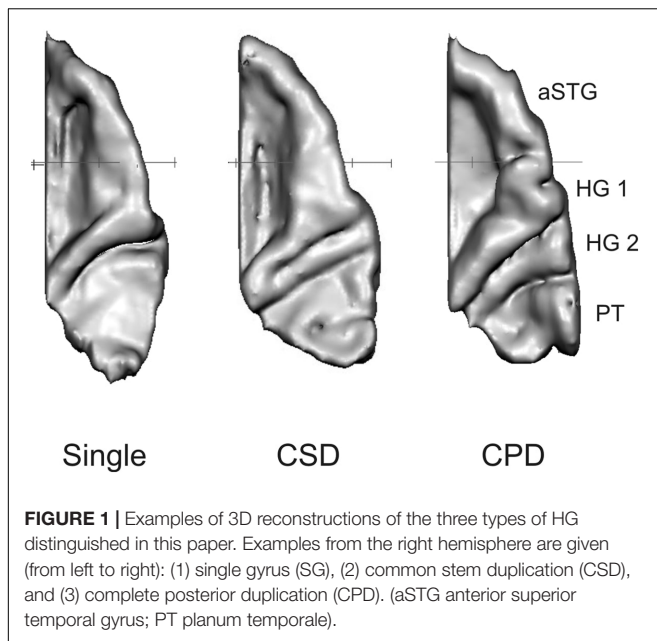
Working Memory Capacity

The importance of working memory capacity for language ability has been shown in various studies and different tasks to measure the components of working memory exist. For this study, three types of tests were applied. Subjects had to do a digit span backward, a digit span forward, and non-word span task. The digit span forward test requires subjects to repeat a rising number of digits (starting out with a small number and always adding one per round) in the same order as heard. The digit span backward, in contrast, requires the repetition of heard digits (same procedure as the aforementioned) backward. In the non-word span task participants are asked to repeat non-word syllables in the same order as heard while paying particular attention to the sounds used in these non-words. All participants were given two chances for the same number of digits/non-words, i.e., if the first attempt of repeating a certain amount of digits failed, the subjects heard a different set of the same amount of digits to repeat. Only if both attempts were incorrect, the test was stopped at that point and no points were given. Per correct series, the subjects received one single point.

Morphometric MRI

For the neuroanatomical analysis, high-resolution T1-weighted structural magnetic MRI (Siemens, Magnetom SonataVision, 1,5 Tesla, software version: syngo MR 2004A, 176 DICOM slices, sagittal orientation, slice thickness 1 mm) were performed in order to investigate the morphology of auditory cortex in both hemispheres. Three-dimensional gray matter surface reconstructions of the auditory cortex (HG) and the planum temporale (PT) were analyzed using a standardized individual approach. This allows for a closer look at the shape of HG in the subjects' brains (Schneider et al., 2002, 2009; Seither-Preisler et al., 2014; Serrallach et al., 2016; Benner et al., 2017). Brain Voyager software QX 2.8 (Brain Innovation B.V, Maastricht, Netherlands) was used for the segmentation of the aforementioned auditory-related areas. Pre-processing steps included the adjustment of brain images in contrast and brightness, as well as a correction for inhomogeneity and a rotation in direction of the antero-posterior commissural line. Normalization in stereotactic space (Talairach and Tournoux, 1988) was carried out to arrive at comparable reconstructions. In the process of segmentation, the superior temporal plane, including HG, the anterior superior temporal gyrus and the planum temporale, were segmented into sagittal MRI slices along the lateral fissure using the standard definition of the landmarks of AC. After this semi-manual slice-by-slice segmentation (adapted from Schneider et al., 2005, 2009; also applied by Wengenroth et al., 2010, 2013; Seither-Preisler et al., 2014; Serrallach et al., 2016; Benner et al., 2017), the auditory cortices of all subjects were 3D-reconstructed and the authors compared the shape of HG in each hemisphere. The three categories chosen for categorization were (1) single gyrus (SG), (2) common stem duplication (CSD) and (3) CPD. In **Figure 1** the three types of HG are compared. These categories are in accordance with

---

TABLE 1 | A description of the different parts of the Modern Language Aptitude Test (MLAT) used in this study (Parts III, IV, and V).

MLAT	Name	Task
Part III	Spelling clues	Sound-symbol association ability and vocabulary knowledge – correct synonyms of disguised words have to be selected (multiple choice).
Part IV	Words in sentences	Grammatical sensitivity – components of sentences have to be identified (grammatical function) and related to elements in other words.
Part V	Paired associates	Associative rote memory – as many words in Kurdish have to be memorized as possible (presented with english translations).

FIGURE 1 | Examples of 3D reconstructions of the three types of HG distinguished in this paper. Examples from the right hemisphere are given (from left to right): (1) single gyrus (SG), (2) common stem duplication (CSD), and (3) complete posterior duplication (CPD). (aSTG anterior superior temporal gyrus; PT planum temporale).

recent research (Marie et al., 2016; Benner et al., 2017) with the exception of multiple gyri, which were only present in four hemispheres of this study and therefore considered to belong to the CSD group in the case of a z-shape ($N = 2$) and to belong to the CPD in the case of more than one CPD ($N = 2$). Lateral HG duplications were considered to be part of the planum temporale and medial duplications (Schneider et al., 2005) to be a sub-form of CSD.

RESULTS

Behavioral Results

Descriptive Results

First of all, a brief summary of the descriptive results of the variety of tests shall be given. In this study, both the English pronunciation score and the Hindi imitation score are considered as measures for language aptitude. Performance in the English pronunciation task was rather high ($M = 6.40$; $SD = 1.72$) in contrast to the Hindi task. In the Hindi speech imitation task, subjects obtained between 2.72 and 7.74 ($M = 4.81$, $SD = 1.64$) of maximally 10 achievable points, with the mean being much lower than in the English task. As also reported by Dogil and Reiterer (2009), the native speakers that had been mixed into the rating procedure additionally received scores from 8 to 10, i.e., the raters considered scores between 8 and 10 as reflecting native performance. The maximum of points achieved by a German-speaking subject was 7.74, which is strikingly high given that the subject had never been exposed to Hindi. According to speech imitation performance in the Hindi test, subjects with a score below 4 were considered to have very poor skills and were classified as 'non-talented,' while subjects with a score above 5 were classified as 'talented'.

The number of instruments subjects had learnt ranged from zero to three ($M = 1.23$, $SD = 0.97$), with most participants playing one single instrument. In stark contrast, the number of foreign languages acquired ranged from one to nine ($M = 2.59$, $SD = 1.72$), although most subjects had learnt two to three foreign languages.

AMMA tonal results ($M = 28.72$, $SD = 5.68$) were quite similar to AMMA rhythm results ($M = 31.10$, $SD = 4.61$). The total score for the AMMA, subsuming both aforementioned parts, ranged from 42 to 79, i.e., it showed a considerable variability ($M = 59.8$, $SD = 10.05$). In the working memory scales, digit span forward ($M = 9.59$, $SD = 1.88$) and digit span backward ($M = 8.76$, $SD = 2.13$) gave similar results. However, subjects performed better in the forward task. With a range from 6 to 14, some participants showed remarkable results, which were far beyond the norm. The digit span backward scores ranged from 4 to 13, which was still higher than the span for the non-word task ($M = 7.55$, $SD = 1.74$), where subjects scored between 5 and 11 points.

Great variability was found in the MLAT total scores with a range from 49 to 109 points (raw score; $M = 83.41$, $SD = 14.23$), reflecting the large gap between 'highly gifted' and 'poor' language learners. The MLAT total score summarizes the results of part III ($M = 36.69$, $SD = 8.62$), part IV ($M = 29.28$, $SD = 5.58$) and part V ($M = 17.31$, $SD = 5.09$). The best performance was thus found in part III, measuring phonetic coding ability, and the least successful performance in part V, the vocabulary learning task.

Correlational Analysis

As the correlational analyses include multiple variables and comparisons, this may increase the risk of chance findings at a critical p-value of 0.05, due to alpha error accumulation. Therefore, the original correlational analysis was complemented by an analysis corrected for multiple comparisons. Similar to the classical Bonferroni correction, the Benjamini–Hochberg procedure we applied is only appropriate for variables independent of each other (McDonald, 2014). Since there is an interdependence between the composite variables AMMA total (consisting of AMMA rhythm and AMMA tonal results), MLAT total (consisting of the three subtests), and an overall pronunciation aptitude score (summarizing the Hindi and English score), this prerequisite was not fulfilled for these variables and they therefore had to be excluded. We used the method of false discovery rate to control for alpha error accumulation in multiple comparisons (Benjamini and Hochberg, 1995). According to the recommendation of the authors, who suggest rates between 0.10 and 0.25 (not to be confounded with regular significance levels, which are much lower), we selected a value of 0.2, corresponding to an acceptable proportion of false discoveries ≤20%. An overview of the correlational results is given in **Table 2**.

AMMA rhythm and AMMA tonal ($r = 0.911$, $p = 0.000$) showed a very strong relationship with each other. Likewise, different parts of the MLAT correlated significantly with each other, namely MLAT IV and MLAT III ($r = 0.590$, $p = 0.001$), but also with the English pronunciation score (MLAT III and English: $r = 0.756$, $p = 0.002$; MLAT IV and English: $r = 0.557$,

TABLE 2 | An overview of the correlational analysis performed with SPSS 22.

	Instruments	Languages	Hindi	English	AMMA tonal	AMMA rhythm	MLAT V	MLAT IV	MLAT III	Non-word	Digit backward	Digit forward	Age
Instruments	1	0.159	0.394***	0.310	0.455***	0.407***	−0.021	0.506***	0.182	0.350	0.282	0.145	0.169
Languages		1	0.003	−0.066	0.218	0.213	0.048	0.217	0.131	0.103	0.273	0.056	0.501***
Hindi			1	0.390***	0.284	0.278*	0.356	0.384*	0.290	0.483***	0.369*	0.447***	0.248
English				1	0.268	0.393***	0.333	0.557***	0.756***	0.163	0.251	0.195	0.000
AMMA tonal					1	0.911***	−0.023	0.333	0.243	0.312	0.260	−0.101	0.168
AMMA rhythm						1	−0.009	0.367	0.365	0.313	0.224	0.034	0.042
MLAT V							1	0.278	0.299	0.012	0.165	0.212	−0.028
MLAT IV								1	0.590***	0.083	0.237	0.246	0.062
MLAT III									1	0.012	0.268	0.320	0.022
Non-word										1	0.335	0.323	0.534***
Digit backward											1	0.259	0.313
Digit forward												1	−0.145
Age													1

***Significant after correction for multiple comparisons (Benjamini–Hochberg procedure). *Significant at p < 0.05 before correction.

$p = 0.002$). MLAT IV and the number of instruments learnt by a subject ($r = 0.506$, $p = 0.005$). Age of participants (ranging from 20 to 40 years) highly correlated with the number of languages ($r = 0.501$, $p = 0.005$) and non-word span ($r = 0.534$, $p = 0.003$). After correction, non-word span and Hindi ($r = 0.482$, $p = 0.008$) and digit span forward and Hindi ($r = 0.447$, $p = 0.015$) still had a strong relationship. The same can be reported for results on Hindi and English tasks ($r = 0.390$, $p = 0.033$), and the Hindi results and number of instruments ($r = 0.394$, $p = 0.032$). Last but not least, AMMA tonal correlated significantly with the number of instruments ($r = 0.455$, $p = 0.013$), whereas AMMA rhythm correlated both with the number of instruments ($r = 0.407$, $p = 0.028$) and the English pronunciation score ($r = 0.393$, $p = 0.035$).

Principal Component Analysis

In order to gain insights into the most influential factors underlying performance on the different test scales, we calculated a principle component analysis (PCA). This included the same scales as shown in the correlation matrix of **Table 2**. A pre-analysis of our data showed that the requirements for the application of the method were fulfilled [(a) the determinant as an indicator of multicollinearity, which should be $p < 0.05$, was $p = 0.002$; (b) the Kaiser-Meyer-Olkin criterion as a measure for the suitability of the sample, which should be above $p = 0.5$, was $p = 0.653$; (c) the Bartlett-test for sphericity, which should be significant at least at $p < 0.05$ was significant at $p < 0.000001$]. We used varimax rotation with Kaiser-normalization, which according to the scree-plot yielded a solution with three factors with eigenvalues clearly above 1. The variance explained by the model was 61.2%, which confirms its appropriateness. **Table 3** shows the rotated component matrix with the coefficients of each scale on the three identified components.

According to the criterion for strong and thus particularly relevant loadings (>0.5), the first component comprises scales related to musicality (AMMA tonal and rhythm scores, number of played instruments), the second component refers to language talent (total English score and parts III, IV, and V of the MLAT) and the third component refers to working memory capacity (digit span forward, digit span backward, non-word span and the Hindi score). Apart from these main findings, there are weaker but still noteworthy loadings (>0.3). These show that the number of instruments played is also associated with the component working memory capacity, while the number of languages spoken is also associated with the component musicality. Moreover, part IV of the MLAT, which measures grammatical sensitivity and is first and foremost a scale of language talent, is also related to musicality. Similarly, non-word span also moderately loads on the component musicality. The total Hindi score, being most strongly related to working memory, also moderately loads on the component language talent.

In a next step, the participants' individual factor scores on each of the three identified components (positive/negative: above/below average; $M = 0$, $SD = 1$) were compared for the three types of right-hemispheric HG morphology (single, common stem, complete duplication). Results are graphically illustrated in **Figure 2**. For all three factors,

TABLE 3 | Rotated component matrix with loading coefficients for each scale.

Rotated component matrix			
	Components		
	1	**2**	**3**
AMMA tonal	**0.918**		
AMMA rhythm	**0.872**		
n° instruments	**0.598**		0.330
n° languages	0.357		
MLAT III		**0.868**	
MLAT IV	0.340	**0.711**	
MLAT V		**0.560**	
English score		**0.829**	
Hindi score		0.312	**0.734**
Non-word span	0.372		**0.729**
Digit span forward			**0.716**
Digit span backward			**0.558**

Extraction method: Principal component analysis. Rotation method: Varimax with Kaiser normalization.

(a) Rotation converged in 6 iterations.

For the sake of clarity, coefficients below 0.3, which signify very small and hence negligible loadings, are not shown. Coefficients above 0.5, which signify particularly relevant contributions, are marked in bold.
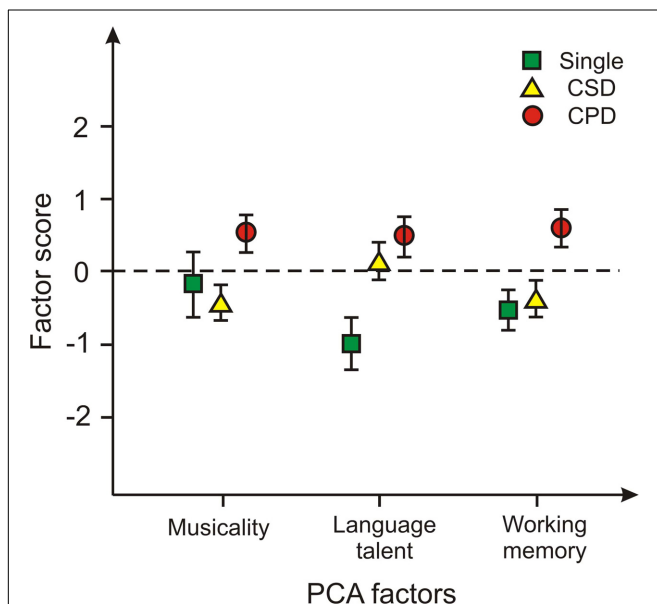

FIGURE 2 | Individual factor scores (positive/negative: above/below average; $M = 0$, $SD = 1$) on each of the three identified PCA components compared for the three types of HG in the right hemisphere. Error bars: SEM (standard error of the mean); Single, single gyrus; CSD, common stem duplication; CPD, complete posterior duplication.

performance of subjects with complete duplications was highest. Concerning musicality, subjects with a complete duplication were significantly better ($M = 0.48$, $SD = 0.97$) than subjects with common stem morphology [$M = -0.4$, $SD = 0.76$; $t_{(20)} = -2.4$, $p = 0.027$]. With regard to language talent, subjects with a

complete duplication were significantly superior ($M = 0.48$, $SD = 0.80$) to subjects with a common stem morphology [$M = 0.15$, $SD = 0.75$; $t_{(16)} = -2.7$, $p = 0.015$] and to subjects with a SG [$M = -0.99$, $SD = 1.0$; $t_{(16)} = -3.4$, $p = 0.004$]. Also with regard to working memory, subjects with a complete duplication were significantly better ($M = 0.65$, $SD = 0.91$) than subjects with a common stem morphology [$M = -3.2$, $SD = 0.86$; $t_{(20)} = -2.6$, $p = 0.019$] and subjects with a SG [$M = -0.52$, $SD = 0.88$; $t_{(16)} = -2.7$, $p = 0.016$].

Hindi Speech Imitation Score

A *t*-test based on the distinction between talented and non-talented subjects revealed significant differences between the two groups. This was the case for the number of played instruments [$t_{(28)} = -2.32$, $p = 0.028$], the English pronunciation score [$t_{(28)} = -2.1$, $p = 0.045$], the digit span forward score [$t_{(27)} = -2.73$, $p = 0.011$], the non-word repetition score [$t_{(27)} = -2.5$, $p = 0.017$] and the MLAT total raw score [$t_{(27)} = -2.27$, $p = 0.032$].

Additionally, a linear multiple regression analysis (method: step-wise forward) was performed for the criterion variable Hindi score and the predictors AMMA tonal, AMMA rhythm, English proficiency, MLAT III, MLAT IV, MLAT V, digit span forward, digit span backward, non-word span, number of instruments, and number of learned languages. The model yielded a corrected R^2-value of 0.375, corresponding to an explained variance of 37.5%, and beta-values (relative importance of contributing variables, summing up to 1) of 0.44 for non-word span, 0.31 for the number of instruments played, and 0.25 for part V of the MLAT. In other words, the three most important predictors for the Hindi speech imitation score were performance on non-word span (i.e., working memory capacity), the number of instruments played by an individual and results of MLAT V, measuring grammatical sensitivity. Overall, these three predictors explain 37,5% of the Hindi score, which points to a high explanatory value of the considered variables for phonetic coding ability.

Results of the Neuroanatomical Analysis

We compared the 3D-reconstructed HGs of both hemispheres in all subjects. First, we categorized all HGs by description (i.e., defining complete duplications and CSDs) as in previous papers (Schneider et al., 2005; Seither-Preisler et al., 2014; Benner et al., 2017). The frequencies of different HG types (altogether $N = 30$) found in our subjects are given in **Table 4**, group-averaged AC

TABLE 4 | Frequency of types of HG in right and left hemispheres in subjects with high and low Hindi score.

		RH (high/low)	LH (high/low)
	Total number (%)	30 (100%)	30 (100%)
Types of HG	Single	8 (26%) (1/7)	18 (60%) (9/9)
	CSD	11 (37%) (3/8)	3 (10%) (2/1)
	CPD	11 (37%) (10/1)	9 (30%) (3/6)

RH, right hemisphere; LH, left hemisphere; CSD, common stem duplication; CPD, complete posterior duplication.

surfaces are presented in **Figure 3** and the individual auditory cortices of all subjects are provided in **Figure 4**.

Figures 3 and **4** nicely portray the differences found in HG morphology in the right hemisphere. It is clearly evident that individuals with high speech imitation aptitude in the Hindi testing, and also individuals with very high scores in the AMMA testing, showed more CPDs of their HG in the right hemisphere. This means that subjects with excellent scores in the language
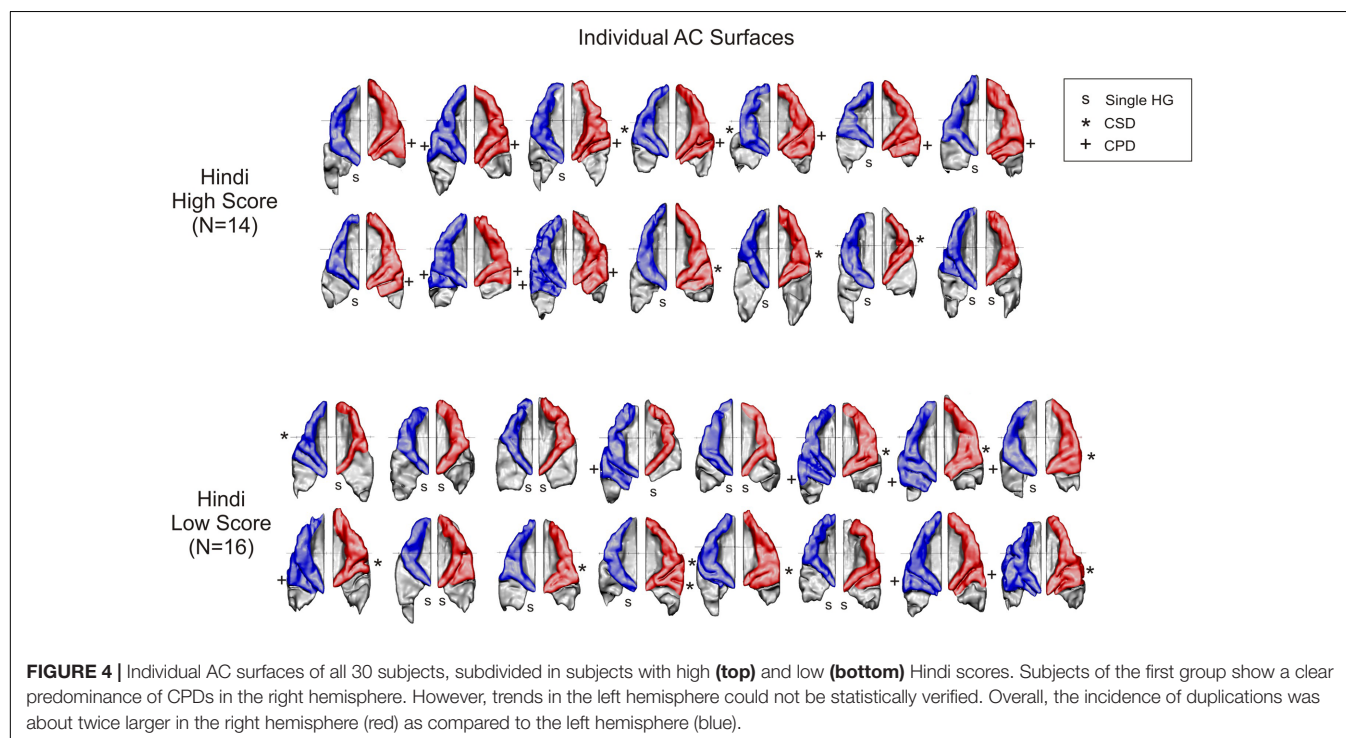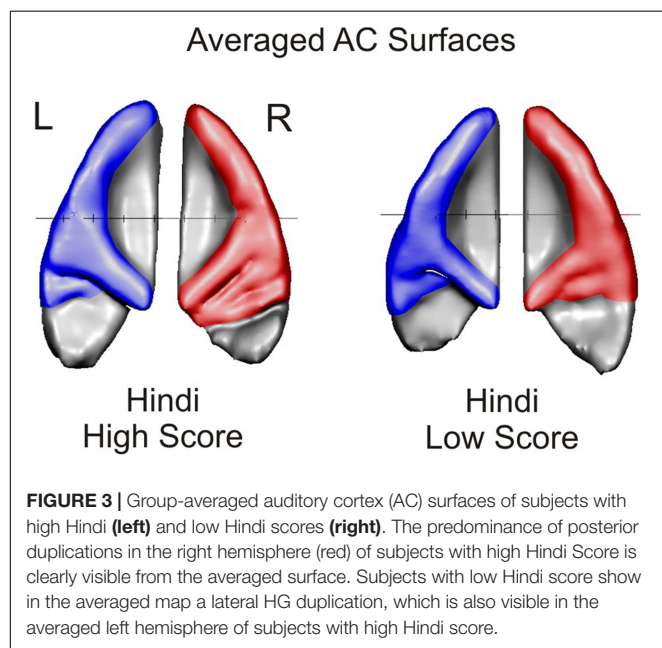
aptitude and in the AMMA testing have two equally prominent HGs in the right hemisphere, in contrast to those with rather low scores, who possess most frequently single gyri or a CSD.
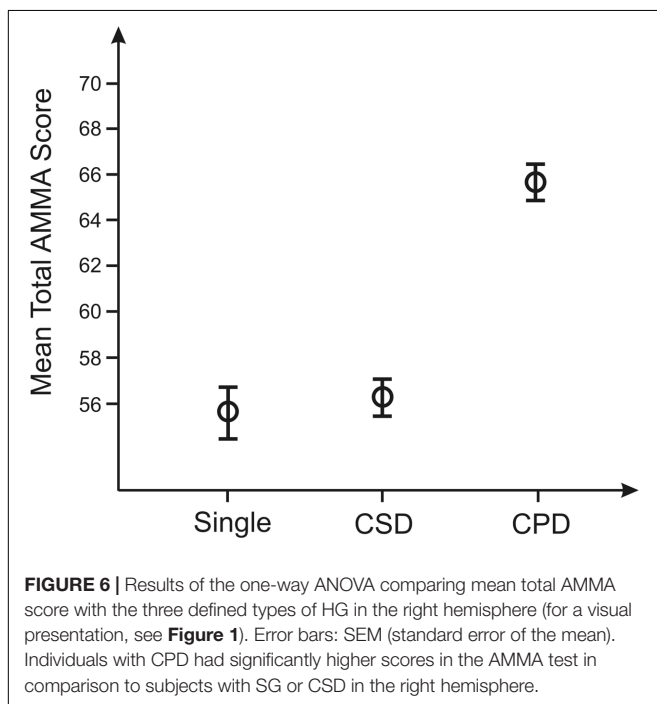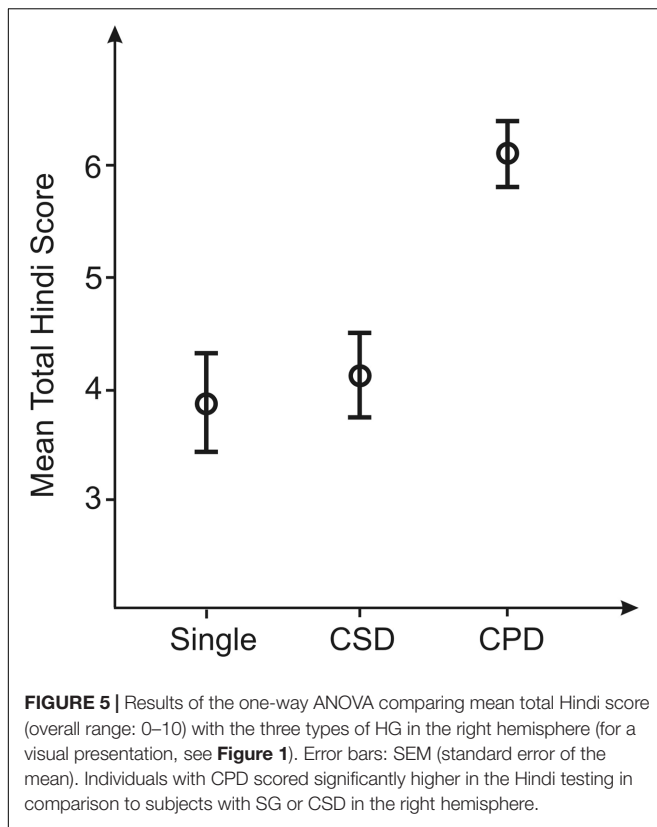
In order to verify the significance of the described exemplary observations, we performed one-way ANOVAs on the Hindi and AMMA test scores for subjects displaying one of the three following morphological HG characteristics in their right hemisphere: (1) SG, (2) CSD, and (3) CPD (double gyrus; CPD). Furthermore, χ^2-tests were performed on the frequency distributions of these neuroanatomical characteristics.

A significant group difference was observed for the Hindi speech imitation score $[F(2,27) = 9.2, p < 0.001, \eta_p^2 = 0.41]$ (**Figure 5**). Subjects with a CPD achieved significantly higher scores (6.1 ± 1.2) than subjects with a SG ($3.9 \pm 1.4; p = 0.002$) and subjects with a CSD ($4.1 \pm 1.4; p = 0.004$). There was no significant difference in Hindi imitation between individuals with a SG and a CSD. This is also reflected by the fact that among the high performers in the Hindi speech imitation task CPD in the right hemisphere occurred most frequently (71%) while in low performers they occurred most rarely [6%; $\chi^2(2) = 14.1, p < 0.001$].

Similar results were found for the AMMA test (**Figure 6**). The mean of the total AMMA score in the right hemisphere for SG was 55.7 ± 3.5, for CSD 56.5 ± 2.8, and for CPD 65.8 ± 2.8. Individuals with CPD achieved significantly higher scores than subjects with SG and CSD $[F(2,26) = 3.8, p = 0.036, \eta_p^2 = 0.23]$. There was no significant difference in the musicality test for SG and CSD.

These findings are also supported by the results of the PCA as presented in **Figure 2**. It is quite evident, that CPD are advantageous for all three components as revealed by the analysis.

FIGURE 3 | Group-averaged auditory cortex (AC) surfaces of subjects with high Hindi **(left)** and low Hindi scores **(right)**. The predominance of posterior duplications in the right hemisphere (red) of subjects with high Hindi Score is clearly visible from the averaged surface. Subjects with low Hindi score show in the averaged map a lateral HG duplication, which is also visible in the averaged left hemisphere of subjects with high Hindi score.

FIGURE 4 | Individual AC surfaces of all 30 subjects, subdivided in subjects with high **(top)** and low **(bottom)** Hindi scores. Subjects of the first group show a clear predominance of CPDs in the right hemisphere. However, trends in the left hemisphere could not be statistically verified. Overall, the incidence of duplications was about twice larger in the right hemisphere (red) as compared to the left hemisphere (blue).

FIGURE 5 | Results of the one-way ANOVA comparing mean total Hindi score (overall range: 0–10) with the three types of HG in the right hemisphere (for a visual presentation, see **Figure 1**). Error bars: SEM (standard error of the mean). Individuals with CPD scored significantly higher in the Hindi testing in comparison to subjects with SG or CSD in the right hemisphere.

FIGURE 6 | Results of the one-way ANOVA comparing mean total AMMA score with the three defined types of HG in the right hemisphere (for a visual presentation, see **Figure 1**). Error bars: SEM (standard error of the mean). Individuals with CPD had significantly higher scores in the AMMA test in comparison to subjects with SG or CSD in the right hemisphere.

In other words, individuals with higher scores in the three components (i.e., musicality, language aptitude and working memory) also had more CPD.

DISCUSSION

The results of the principal component analysis, performed to gain insights into the most influential factors underlying performance on test scales, revealed three clear components, which very nicely reflect the three core aspects investigated in this study. They are (1) musicality (AMMA tonal, AMMA rhythm, number of instruments), (2) language aptitude (English score, all parts of the MLAT) and (3) working memory capacity (including the Hindi score and all working memory tests) (see **Table 3**). Since the main aim of our research was to explore the connection between these three variables and especially their relationship with auditory cortex morphology, they will be discussed separately in the upcoming paragraphs.

Language Aptitude

Language aptitude is at the heart of our research and deserves sufficient attention with this regard. The Hindi score is the variable we assume to measure phonetic coding ability, i.e., a measure of a subcomponent of language aptitude. In the PCA, the Hindi score loaded on two components, namely working memory capacity and language aptitude. This nicely reflects the fact that working memory is an essential aspect of language aptitude and that the Hindi score can be seen as an indicator of both. Generally, the results of the PCA were very clear with regard to the component of language aptitude, which was clearly dominated by the MLAT, the English score and the Hindi score. This strongly supports their validity for measuring language aptitude.

The Hindi speech imitation task and the English pronunciation task measure very different components of language learning ability. The Hindi score is a speech imitation score, which requires the reproduction of unknown speech material. Still, it also demands the accurate auditory processing of this material. Otherwise, no successful imitation can take place. The English pronunciation task, in contrast, gives an overview of a subject's pronunciation skills in their second language. Whereas the first is supposed to be a measure of phonetic coding ability (a major component of language aptitude), the second measures pronunciation proficiency in an already acquired language. High pronunciation proficiency, however, relies on a certain ability for phonetic coding and the two scores for language learning ability hence clearly go hand in hand. Even after correction for multiple comparisons, a positive, moderate correlation could be found between the two scores ($r = 0.395$, $p = 0.031$). The Hindi score could be seen as a precursor for the English score since high phonetic coding ability should lead to an excellent pronunciation in any language acquired by an individual. An issue with English in this case might be that many of the subjects had spent considerable time in an English-speaking country or had even studied English. English is a lingua franca and as education in Germany introduces English as first foreign language for every child, acquiring a native-like pronunciation is already supported from the beginning. Moreover, children are exposed to English in the early years of their lives, which might influence the success of their acquisition (critical periods are not well-defined but assumed to exist). This could explain why even individuals

with lower scores in the Hindi test had a good pronunciation in English. We argue that frequent and long-lasting language exposure and contact are certainly reasons why the English scores were significantly higher. Since we see language aptitude as a predominantly innate capacity that unfolds over time in interaction with the environment, we assume that it is rather the Hindi score that should predict the English score, i.e., the better individuals are at decoding, retaining and reproducing unknown speech material, the easier it should be for them to develop excellent pronunciation skills in a given language.

Moreover, clear significant positive correlations were found between the English score and MLAT parts III and IV. The results in these two cases were quite robust, leading to two possible interpretations. First, the English score is related to an individual's English skills and as the MLAT had to be used in English, it was probably more the subjects' English proficiency that led to the high results. Second, high pronunciation proficiency in English is the result of particularly high language aptitude (phonetic coding ability and grammatical ability), as highlighted by the results of the correlational analysis. If we assume that the subtests of the MLAT, test III measuring phonetic ability and test IV measuring grammatical sensitivity, are excellent indicators of these components of language aptitude, we would at least expect a very high correlation between the Hindi speech imitation score and MLAT part III, which was not the case. But, as already mentioned, one of the major problems with the MLAT is that it is in English, giving individuals with better English skills a clear advantage over those subjects with less proficiency in English. Unfortunately, no German version of the MLAT exists to date, but in the past years language-independent tests, such as the LLAMA language aptitude battery (Meara, 2005), have gained popularity (Granena, 2013; Artieda and Muñoz, 2016; Rogers et al., 2016; Kepinska et al., 2017a,b). Therefore, only language-independent tests can exclude a possible influence of language experience and shall be used for future research. To get back to one aforementioned finding, English pronunciation skills also correlated with grammatical sensitivity (MLAT IV) and we assume that this should be rather a result of aptitude than of proficiency. Language analytic ability, the umbrella term under which grammatical sensitivity is nowadays subsumed, is an important component of language aptitude (Kepinska et al., 2017a,b) and should thus be of high significance for the learning of foreign languages. In our case, however, we focused on foreign language pronunciation ability (Jilka, 2009) and can only marginally address the significance of grammatical sensitivity. Looking at the last part of the MLAT, we see that vocabulary learning (part V – associative memory) did not correlate at all with the English score. It seems that vocabulary learning does not depend on proficiency but possibly on other factors such as learning strategies and motivation. Additionally, the claims that verbal working memory skills (attributed to the phonological loop) are essential for novel vocabulary learning (Gathercole and Baddeley, 1990; Atkins and Baddeley, 1998; Baddeley et al., 1998; Gathercole et al., 1999) could not be confirmed in our study, since MLAT V could not be linked to any working memory task.

One result that needs further discussion is the fact that the number of languages spoken by a subject did not show any relationship to any other score, except for age. It would be logical to assume that the more languages you speak, the better you are at learning different aspects of a new language (MLAT) or the better your English pronunciation and your speech imitation skills are. Vice versa, we would expect individuals with very high language learning ability to learn more languages due to the facility with which they acquire foreign languages. This was not the case due to various possible reasons. First, not everybody is willing to learn numerous languages for a number of reasons, e.g., a lack of time, opportunity or necessity. Secondly, having learnt a considerable number of foreign languages does not state anything about a subject's proficiency or their learning process. The sample was limited but the number of foreign languages spoken by the individuals ranged from one to nine, which is quite an outstanding number. Another issue, however, is the fact that we were not able to control in any form how well the participants had learnt the languages and how well they were able to speak them at the time of the testing. It would have been necessary to include proficiency measures (grammar, vocabulary, and pronunciation) in all foreign languages in order to find the specific reasons why no correlations could be found. We further conclude that this result, i.e., that the number of languages does not impact an individual's language aptitude, strongly supports our claim that language aptitude is a rather innate and inflexible capacity that cannot be altered through learning or practice at least. Although Eisenstein (1980) found that previous language training and bilingualism led to higher scores on the MLAT and also Thompson (2013) claimed that previous language experience may alter language aptitude, our study could not corroborate these findings. In the first case, this could also be explained due to excellent English skills. Since we assume, as already mentioned, that language aptitude is a trait somehow present before the acquisition of any language, it should not make a great difference whether an individual had learnt to speak three or nine languages. Earlier, language aptitude was seen as a stable construct that cannot be modified or developed through practice (Carroll, 1958 among others). Even though these assumptions have been questioned in the past decades (Klein, 1995; Sáfár and Kormos, 2008; Thompson, 2013), aptitude may not be as dynamic a construct as claimed by some researchers.

Other variables correlating positively and significantly with the Hindi speech imitation score were the number of instruments played by a subject and two tests of working memory capacity, namely digit span forward and non-word repetition. As has been discussed quite extensively in the introduction section, musical ability is very important for foreign language learning and playing an instrument certainly enhances auditory processing in an individual. For this reason, we expected a strong relationship between musicality scores (AMMA and number of instruments) with the Hindi speech imitation test (for details, see next paragraphs).

Musicality

First, the PCA clearly defined three most influential factors for the musicality component, namely our musicality test (AMMA)

and the number of instruments a subject played. It is noteworthy, however, that three other variables load on this component. The observation that non-word span also loads on the component musicality probably reflects the fact that musical processing builds upon temporally structured, but semantically undefined information that has to be kept in working memory. On the other hand, the fact that MLAT IV, which measures grammatical sensitivity, is moderately related to musicality, is only slightly surprising, as the understanding of language and music depends on internalized grammatical rules. It seems that the PCA factor musicality had a higher validity than the single subscales alone, which is evident from the fact that the AMMA scores and the number of instruments played were only positively but not significantly correlated with the number of languages spoken by subjects (see **Table 2**). The number of languages might thus be related to musicality on a more general level, which could not be sufficiently captured by the single tests.

Concerning musicality, the number of instruments played by an individual has often been assumed to have a considerable impact on a variety of cognitive skills, language acquisition just being one of them (Milovanov et al., 2008; Nardo and Reiterer, 2009). Music has very accurately been described as a resource that leads to auditory fitness (Kraus and Chandrasekaran, 2010) and positively influences the acquisition of skills in other domains, a phenomenon termed positive transfer. The correlations found in this study show a strong relationship between the musical domain and several language-relevant skills. Those who had very good scores in the language aptitude tests, also played more instruments and had better scores on the musicality tests. This supports findings from very recent research and confirms the strong relationship between the two (Milovanov et al., 2004, 2008, 2009, 2010; Magne et al., 2006; Dogil and Reiterer, 2009; Fonseca-Mora et al., 2011; Christiner and Reiterer, 2013; Seither-Preisler et al., 2014; Lee and Lin, 2015; Schön et al., 2004 among others).

Moreover, the number of instruments played by an individual also correlated positively and significantly with the Hindi score, which further confirms the expected strong relationship between music and language (Milovanov et al., 2008; Christiner and Reiterer, 2013; Seither-Preisler et al., 2014; Serrallach et al., 2016). In addition, the two parts of the AMMA correlated significantly with each other supporting the fact that people who have a certain musical ability are very good in different musical domains, in this case rhythmic and melodic discrimination abilities. The two AMMA parts further correlated positively and significantly with the number of instruments played showing that individuals who learn to play more instruments also have better auditory discrimination abilities, i.e., a functionally more efficient auditory cortex (Kraus and Chandrasekaran, 2010). The moderate correlation between the Hindi score and the AMMA test can be explained by the simple fact that the amount of time subjects had played the instruments and the amount of practice they had put into the learning process were not taken into account. These are definitely factors that need to be taken into consideration in future research.

Another interesting finding was the strong correlation between AMMA rhythm and English pronunciation skills. This was quite unexpected, in particular because no relationship with the Hindi score could be found. One option is that high rhythm skills and good rhythm perception facilitate the acquisition of a language and due to considerable practice and experience over time, this improves the participants' pronunciation. In addition, the text used in this study, namely 'The North Wind and the Sun,' is very lyric-like (a fable for children, to be more precise) and may thus be associated with rhythmic perception or give an advantage to musically gifted individuals.

To conclude, for future research it will be important to spend more time investigating the concept of musical aptitude or musicality and using a variety of measures with the aim of fully grasping the construct. There are surely more factors that need to be taken into account and although most studies in this area use the AMMA test as a standard measure for musicality, it would be useful to additionally calculate an index of musical practice (see Seither-Preisler et al., 2014; Serrallach et al., 2016). This provides a fine-grained measure for musical expertise and it allows the implementation of numerous aspects of musicality (different music-related skills and associated variables such as amount of practice, singing interest etc.). This will surely be of high relevance when further investigating the relationship between language and music.

Working Memory Capacity

It does not come as a surprise that speech imitation requires excellent working memory skills and the claim that working memory makes up quite a considerable part of language aptitude is surely not far-fetched (Miyake and Friedman, 1998; Wen and Skehan, 2011; Wen, 2012; Wen et al., 2017). Other studies that have challenged this assumption found that speech imitation skills rely heavily on working memory (Ellis and Sinclair, 1996; Miyake and Friedman, 1998; Kormos and Sáfár, 2008; Sáfár and Kormos, 2008; Biedroń, 2012; Linck et al., 2013). Consistently, we found a positive relationship between speech imitation skills and different measures of working memory, as indicated in all studies above. The three measures we applied were digit span forward, backward, and non-word span. Two of them correlated positively and significantly with the Hindi score, the non-word span showing the highest correlation with this respect. It was only the Hindi test and the age of participants, that showed a strong relationship to non-word span, however. It is common knowledge that children cannot be compared to adults with regard to measures of working memory since working memory seems to improve over time. In this study, this could only be confirmed for non-word span and not for the other two working memory tasks. It seems questionable that working memory capacity improves with age in individuals between 20 and 40 years. Rather, we assume that it might be that the older participants were also those with generally better non-word span, which led to the finding.

In the correlational analysis, the three working memory scores could not be linked to any other variable. The PCA, in contrast, showed a clearly defined component for working memory capacity, based on test results in the Hindi testing, non-word span, digit span forward and backward. In this analysis, the number of instruments loaded on this component as well, supporting a possible influence between musical expertise

and high working memory capacity. Moreover, the PCA also points toward a stronger relationship between the three working memory scores, and also with the Hindi score, than shown in the correlational analysis.

One of the main assumptions of this study was that phonetic coding ability is the component of language aptitude that should be best measured through the Hindi test. However, it seems that the non-word span captures a very similar ability. Furthermore, the non-word test has also been used as an indicator for specific language impairment (Botting and Conti-Ramsden, 2001; Coady and Evans, 2008), supporting our hypothesis of it being a test measuring high or particularly low language learning ability. The Hindi test requires the decoding of unfamiliar speech, retaining it for a particular amount of time and the ability to reproduce it as correctly as possible. Despite some slight differences, both the non-word task and the Hindi speech imitation task use speech material that basically consists of simple CV (consonant-vowel) syllables. So whereas the working memory load increases in the non-word span, the same syllables have to be repeated only adding one element at a time. In the Hindi task, the words and sentences always change and something completely new has to be reproduced. In sum, both heavily rely on working memory capacity and we propose that both tests are equally significant and useful measures of working memory capacity on the one hand, and phonetic coding ability on the other hand. We therefore propose to further develop non-word tests in order to ameliorate language aptitude testing batteries (Chan et al., 2011).

Digit span forward also correlated positively and significantly with the Hindi score, but not with any other score. This is a little surprising because other studies (Van den Noort et al., 2006; Kormos and Sáfár, 2008; Biedroń, 2012) have shown that high ability and success in foreign language acquisition, in our case the English score, correlate with working memory tasks of differing complexity. Nevertheless, considering that the Hindi score is our main language aptitude score, we conclude that both simple and complex working memory skills are required to imitate foreign speech, i.e., for phonetic coding ability. This further supports the hypothesis that working memory is a core component of foreign language aptitude (Wen, 2016; Wen et al., 2017) but we do not agree with the hypothesis that working memory may be seen as an equivalent to language aptitude (Linck et al., 2014).

Last but not least, one surprising finding is a lack of relationship between the three working memory tests. The construct of working memory includes different components, which are expected to influence one another or at least share some common basis. Our study, however, did not show any correlation between the three scores. Only the PCA showed the dependence between the three variables and also the Hindi testing. We therefore propose that different components of working memory are indeed relevant for language learning but to a certain extent independent from each other.

Neuroanatomic Markers for Language Aptitude and Musicality

Studies of the past years have partly investigated the neural basis of language learning ability and they have certainly highlighted the significance of the structure of language-relevant regions in the human brain. As Berken et al. (2015) correctly summarizes, structural variation in the brain can indeed reveal variation in language aptitude. Learning novel elements of a language, e.g., tonal pitch contrasts and phonetic differences (Golestani and Zatorre, 2009), and perceiving and producing novel speech sounds (Golestani et al., 2002; Golestani and Pallier, 2007) can reveal interesting information as to which regions are important for these processes. Mostly, language has been ascribed to the left hemisphere and also findings regarding HG (with respect to language) have emphasized the special role of the left side (Golestani and Pallier, 2007; Golestani et al., 2007; see Introduction).

The most interesting finding of the neuroanatomic analysis of this study is that individuals with high Hindi scores also had more CPDs of HG in the right hemisphere, contradicting theories of leftward lateralization for language functions in healthy adults. In this regard it is revealing that also the AMMA score showed a particularly strong relationship with HG duplications in the right hemisphere. Evidently, both skills, even if not being directly correlated with each other, appear to be closely linked to right HG. Most interestingly, when the participants' individual factor scores on each of the three components were compared for the three HG types, it became obvious that a CPD seems to be correlated with high results in all three components of the PCA (see **Figure 2**). In other words, for all three factors defined (musicality, working memory, and language aptitude), performance of subjects with CPD was highest.

There are two main topics that need to be discussed accordingly with respect to these findings. First of all, the results suggest more than just a positive relationship between language aptitude, music and partly also working memory. It is necessary to specify the nature of this relationship, the influence of the auditory cortex on the two and why it is only the right hemisphere that seems to be much more important. Second, the function of CPDs in HGs is far from being clear. Leaving aside the hemispheric differences, there is convincing evidence for a specific structure-function relationship of HG duplications and furthermore a considerably larger prevalence of HG duplications in both musicians (Benner et al., 2017) and linguistically talented subjects (Golestani et al., 2011). The connectivity between the first HG, hosting in most cases the primary auditory core areas, and the different HG duplications, hosting among others associated language-related belt and parabelt areas, may have a hitherto unknown impact on auditory functioning and thus the development of language and musical skills.

Since our main aim was to find the neuroanatomical markers of language aptitude and language has been claimed to be predominantly left-lateralized, we expected to find differences mainly in the left hemisphere (Golestani et al., 2002, 2007, 2011; Golestani and Pallier, 2007; Wong et al., 2007; Dogil and Reiterer, 2009; Golestani and Zatorre, 2009; Reiterer et al., 2011; Warrier et al., 2012; Hu et al., 2013). Yet, in the current study the occurrence of duplications in the left hemisphere was considerably lower as compared to the balanced distribution in

the right hemisphere (see **Table 3**). Given the little variation found in the left hemisphere, a much larger sample would have been needed to perform corresponding group statistical comparisons for the left hemisphere. Therefore, all further discussion focuses on the right hemisphere only.

It is well-known that musical ability heavily relies on the right hemisphere and also recent research has shown the significance of HG on the right side for musical processing. But why does HG in the right hemisphere relate so well with the Hindi testing then? There are numerous possible explanations. To put it simple, the results clearly indicate that the shape of auditory cortex and the number of HGs in the right hemisphere are linked to musical and linguistic ability. Relationships with working memory, in contrast, could only be found through the PCA, in which the component of working memory capacity included the Hindi score. Therefore, working memory and AC structure will not be discussed in detail here. To get back to the main issue, only the individuals with a CPD (i.e., two complete gyri) in the right hemisphere had significantly higher scores in the AMMA test and in the Hindi task. The individuals with single gyri and CSDs (not counted as two complete gyri) had substantially lower results in both. Is it therefore necessary to have two HGs in the right hemisphere to have a considerable advantage in auditory processing? And if we assume that individuals with a double HG have a functionally advantageous auditory cortex, why is it that both language and music seem to be so heavily influenced by it?

In a larger sample we might have discovered even more robust evidence for the observed relationship between language aptitude and musicality. Other studies, however, have already shown that musical ability facilitates language learning. The Hindi speech imitation task is basically a working memory capacity task that requires good use of the articulators to produce foreign speech material and a functionally efficient auditory cortex to hear the subtle differences in the speech input. Could it be that only phonetic coding ability, i.e., only this particular component of language aptitude, is highly dependent on (1) musical ability or (2) the processing of auditory cortex in the right hemisphere? If phonetic coding ability were dependent on auditory processing of music-relevant features in the right hemisphere, this would explain why only the right hemisphere showed double gyri in most subjects. One of the most difficult questions in this regard is to what extent differences in auditory cortex are due to language aptitude or due to musical ability. It could be that, as in other studies, we just found a confirmation of the importance of CPDs in the right hemisphere for musical processing and due to the fact that Hindi speech imitation requires non-speech processing expertise, we found a similar relationship between the two. Another possibility is that we found a neuroanatomical marker for foreign language pronunciation aptitude and this marker also influences musical processing, leading to a high capacity in both domains. Third, our results could suggest that the right hemisphere is more important than assumed for elementary auditory processing, which is at the basis of both speech and music. Even though we do not doubt that numerous areas in the brain are of high importance for the processing of language, our results clearly highlight the

significance of auditory cortex as an essential area of auditory processing.

It could be that AMMA and Hindi, which did not show a direct correlation in this study, are independently linked to right hemispheric functions that require more HGs for efficient auditory processing. Other studies have shown that the individual morphology of these structures, despite high inter-individual variation, are extremely stable from childhood into adulthood (Seither-Preisler et al., 2014). It could thus be assumed that individual differences are first and foremost not due to environmental influences or practicing behavior. Rather, they appear to have a strong biological component, which may be genetic, prenatal, or very early post-natal. As yet, it is also unclear of how the gross-morphological structural characteristics of auditory cortex are related to characteristic functional activation patterns. This important aspect should be specifically addressed in subsequent investigations. In particular, the kind of advantage a CPD of HG has in an individual's brain and if and how this affects language learning and musical ability remains to be uncovered. In addition, even if we know that traits such as shape and number of gyri in auditory cortex play a certain role, it cannot be deduced what kind of advantage CPDs give an individual for language or music processing.

We are aware that the view of language aptitude has changed in the past decades and it is more and more frequently referred to as a dynamic construct that may indeed undergo change over time. Quite recently, an appealing study by Kepinska et al. (2017a,b) on language analytic ability highlighted the significance of the right hemisphere for language aptitude. Moreover, Prat et al. (2016) reported right-hemispheric involvement among highly successful L2 learners in their resting-state qEEG study with adult learners. The right hemisphere might thus be more important for language acquisition and processing than initially assumed. More research will be needed to explore the involvement of the right hemisphere, in particular the right HG, in different aspects of language aptitude. Also, given the various regions in the brain that are essential for language processing, we will aim at developing methods in order to structurally analyze other significant areas, such as the inferior parietal lobule (Dogil and Reiterer, 2009; Reiterer et al., 2011; Hu et al., 2013; Golestani et al., 2002, 2007, 2011 among others).

To sum up – if it is possible to determine neuroanatomical markers that remain highly stable from early infancy to adulthood, this challenges the assumption that the capacity to acquire associated behavioral skills can be substantially altered throughout lifetime. Furthermore, if the structures of certain brain regions are strongly related to specific behavioral skills, we have to find out how they control the natural unfolding and development of these skills. Although there is no doubt that numerous external variables also influence the development of language and musical skills, we support the claim that there are strong innate and/or prenatally determined neurological factors that remain to be uncovered in the next decades. We are already working on similar investigations in children with differing degrees of musicality and language ability in order to

confirm and extend the results of this study. We would also like to encourage other researchers to investigate language aptitude from an anatomical viewpoint, additionally to functional differences that have been repeatedly found in individuals with high and low language learning ability.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of 'Ethikkommission des Universitätsklinikums Tübingen (Germany)' with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the 'Ethikkommission des Universitätsklinikums Tübingen.'

AUTHOR CONTRIBUTIONS

All authors contributed to the submitted paper. ST was responsible for the design of the work, data analysis and interpretation and drafting the article. SR contributed extensively to the data collection. All authors, PS, AS-P, SR, and ST performed a critical revision of the paper

and gave their approvement to the final version for submission.

REFERENCES

Abrahamsson, N., and Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Stud. Second Lang. Acquis.* 30, 481–509. doi: 10.1017/S027226310808073X

Al-Shabatat, A. M. (2013). A review of the contemporary concepts of giftedness and talent. *Int. Interdiscip. J. Educ.* 2, 1336–1346. doi: 10.12816/0002983

Artieda, G., and Muñoz, C. (2016). The LLAMA tests and the underlying structure of language aptitude at two levels of foreign language proficiency. *Learn. Individ. Diff.* 50, 42–48. doi: 10.1016/j.lindif.2016.06.023

Atkins, P. W., and Baddeley, A. D. (1998). Working memory and distributed vocabulary learning. *Appl. Psycholinguist.* 19, 537–552. doi: 10.1017/S0142716400010353

Baddeley, A. (2003a). Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4, 829–839. doi: 10.1038/nrn1201

Baddeley, A. (2003b). Working memory and language: an overview. *J. Commun. Disord.* 36, 189–208. doi: 10.1016/S0021-9924(03)00019-4

Baddeley, A. (2017). Modularity, working memory and language acquisition. *Second Lang. Res.* 33, 299–311. doi: 10.1177/0267658317709852

Baddeley, A., Gathercole, S., and Papagno, C. (1998). The phonological loop as a language learning device. *Psychol. Rev.* 105, 158–173. doi: 10.1037/0033-295X.105.1.158

Baddeley, A. D., and Hitch, G. J. (1974). Working memory. *Psychol. Learn. Motiv.* 8, 47–90. doi: 10.1016/S0079-7421(08)60452-1

Baddeley, A. D., and Hitch, G. J. (2000). Development of working memory: should the Pascual-Leone and the Baddeley and Hitch models be merged? *J. Exp. Child Psychol.* 77, 128–137. doi: 10.1006/jecp.2000.2592

Bear, M. F., Connors, B. W., and Paradiso, M. A. (2006). *Neuroscience: Exploring the Brain.* Baltimore, MD: Lippincott Williams & Wilkins.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 57, 289–300.

Benner, J., Reinhardt, J., Schneider, P., Stippich, C., Wengenroth, M., and Blatow, M. (2017). Prevalence and function of Heschl's gyrus morphotypes

in musicians. *Brain Struct. Funct.* 222, 3587–3603. doi: 10.1007/s00429-017-1419-x

Berken, J. A., Gracco, V. L., Chen, J. K., Watkins, K. E., Baum, S., Callahan, M., et al. (2015). Neural activation in speech production and reading aloud in native and non-native languages. *Neuroimage* 112, 208–217. doi: 10.1016/j.neuroimage.2015.03.016

Besson, M., and Schön, D. (2003). "Comparison between language and music," in *The Cognitive Neuroscience of Music*, eds I. Peretz and R. Zatorre (Oxford: Oxford University Press), 269–293. doi: 10.1093/acprof:oso/9780198525202.003.0018

Besson, M., Schön, D., Moreno, S., Santos, A., and Magne, C. (2007). Influence of musical expertise and musical training on pitch processing in music and language. *Restor. Neurol. Neurosci.* 25, 399–410. doi: 10.1162/jocn.2010.21585

Biedroń, A. (2011a). Personality factors as predictors of foreign language aptitude. *Stud. Second Lang. Learn. Teach.* 1, 467–489. doi: 10.14746/ssllt.2011.1.4.2

Biedroń, A. (2011b). "Intelligence in Gifted L2 Learners," in *Second Language Learning and Teaching*, ed. M. Pawlak (Berlin: Springer), 129–142.

Biedroń, A. (2012). "Memory abilities in gifted foreign language learners," in *New Perspectives on Individual Differences in Language Learning and Teaching*, ed. M. Pawlak (Berlin: Springer), 77–96.

Biedroń, A. (2015). Neurology of foreign language aptitude. *Stud. Second Lang. Learn. Teach.* 5, 13–40. doi: 10.14746/ssllt.2015.5.1.2

Biedroń, A., and Pawlak, M. (2016a). New conceptualizations of linguistic giftedness. *Lang. Teach.* 49, 151–185. doi: 10.1017/S0261444815000439

Biedroń, A., and Pawlak, M. (2016b). The interface between research on individual difference variables and teaching practice: the case of cognitive factors and personality. *Stud. Second Lang. Learn. Teach.* 6:3. doi: 10.14746/ssllt.2016.6.3.3

Biedroń, A., and Szczepaniak, A. (2009). The cognitive profile of a talented foreign language learner. A case study. *Psychol. Lang. Commun.* 13, 53–71. doi: 10.2478/V10057-009-0004-7

Botting, N., and Conti-Ramsden, G. (2001). Non-word repetition and language development in children with specific language impairment (SLI). *Int. J. Lang. Commun. Disord.* 36, 421–432. doi: 10.1080/13682820110074971

Brown, H. D. (2006). *Principles of Language Learning and Teaching*. White Plains, NY: Pearson Education.

Carroll, J. B. (1958). A factor analysis of two foreign language aptitude batteries. *J. Gen. Psychol.* 59, 3–19. doi: 10.1080/00221309.1958.9710168

Carroll, J. B. (1962). "The prediction of success in intensive foreign language training," in *Training Research and Education*, ed. R. Glaser (Pittsburgh: University of Pittsburgh Press), 87–136.

Carroll, J. B. (1973). Implications of aptitude test research and psycholinguistic theory for foreign language teaching. *Linguistics* 11, 5–14. doi: 10.1515/ling.1973.11.112.5

Carroll, J. B. (1990). "Cognitive abilities in foreign language aptitude," in *Language Aptitude Reconsidered*, eds T. S. Parry and C. W. Stansfield (Englewood Cliffs, NJ: Prentice Hall), 11–30.

Carroll, J. B., and Sapon, S. (1957). *Modern Language Aptitude Test*. Bethesda, MD: Second Language Testing.

Chan, E., Skehan, P., and Gong, G. (2011). Working memory, phonemic coding ability and foreign language aptitude: potential for construction of specific language aptitude tests–the case of Cantonese. *Ilha Desterro* 60, 45–69.

Chobert, J., and Besson, M. (2013). Musical expertise and second language learning. *Brain Sci.* 3, 923–940. doi: 10.3390/brainsci3020923

Christiner, M., and Reiterer, S. M. (2013). Song and speech: examining the link between singing talent and speech imitation ability. *Front. Psychol.* 4:874. doi: 10.3389/fpsyg.2013.00874

Christiner, M., and Reiterer, S. M. (2015). A mozart is not a pavarotti: singers outperform instrumentalists on foreign accent imitation. *Front. Hum. Neurosci.* 9:482. doi: 10.3389/fnhum.2015.00482

Coady, J. A., and Evans, J. L. (2008). Uses and interpretations of non-word repetition tasks in children with and without specific language impairments (SLI). *Int. J. Lang. Commun. Disord.* 43, 1–40. doi: 10.1080/13682820601116485

Deiglmayr, A., Schalk, L., and Stern, E. (2017). "Begabung, intelligenz, talent, Wissen, kompetenz und expertise: eine Begriffserklärung," in *Begabungen und Talente*, eds U. Trautwein and M. Hasselhorn (Göttingen: Hogrefe Verlag), 1–16.

Dogil, G., and Reiterer, S. M. (2009). *Talent and Brain Activity*. Berlin: Mouton de Gruyter.

Dolman, M., and Spring, R. (2014). To what extent does musical aptitude influence foreign language pronunciation skills? A multi-factorial analysis of Japanese learners of English. *World J. Engl. Lang.* 4, 1–11. doi: 10.5430/wjel.v4n4p1

Dörnyei, Z. (1998). Motivation in second and foreign language learning. *Lang. Teach.* 31, 117–135. doi: 10.1017/S026144480001315X

Dörnyei, Z. (2006). Individual differences in second language acquisition. *AILA Rev.* 19, 42–68.

Dörnyei, Z., and Ryan, S. (2015). *The Psychology of the Language Learner Revisited*. New York, NY: Routledge.

Dörnyei, Z., and Skehan, P. (2003). "Individual differences in second language learning," in *The Handbook of Second Language Acquisition*, eds M. Long and C. J. Doughty (Oxford: Blackwell), 589–630.

Eisenstein, M. (1980). Childhood bilingualism and adult language learning aptitude. *Int. Rev. Appl. Psychol.* 29, 159–174. doi: 10.1111/j.1464-0597.1980.tb00888.x

Ellis, N. C., and Sinclair, S. G. (1996). Working memory in the acquisition of vocabulary and syntax: putting language in good order. *Q. J. Exp. Psychol.* 49, 234–250. doi: 10.1080/027249896392883

Faulkner, H. (2003). *Predicting Gifted Foreign Language Learning and Performance*. Ph.D. dissertation, University of Nottingham, Nottingham.

Fonseca-Mora, C., Toscano-Fuentes, C., and Wermke, K. (2011). Melodies that help: the relation between language aptitude and musical intelligence. *Int. J. Engl. Stud.* 22, 101–118.

Friederici, A. D. (2009). Pathways to language: fiber tracts in the human brain. *Trends Cogn. Sci.* 13, 175–181. doi: 10.1016/j.tics.2009.01.001

Gagné, F. (1995). From giftedness to talent: a developmental model and its impact on the language of the field. *Roeper Rev.* 18, 103–111. doi: 10.1080/02783199509553709

Gagné, F. (2005). "From gifts to talents: the DMGT as a developmental model," in *Conceptions of Giftedness*, eds J. E. Davidson and R. J. Sternberg (Cambridge: Cambridge University Press), 98–119.

Galaburda, A. M., Sanides, F., and Geschwind, N. (1978). Cytoarchitectonic left-right asymmetries in the temporal speech region. *Arch. Neurol.* 35, 812–817. doi: 10.1001/archneur.1978.00500360036007

Ganschow, L., and Sparks, R. (1995). Effects of direct instruction in Spanish phonology on the native-language skills and foreign-language aptitude of at-risk foreign-language learners. *J. Learn. Disabil.* 28, 107–120. doi: 10.1177/002221949502800205

Gathercole, S. E., and Baddeley, A. D. (1990). The role of phonological memory in vocabulary acquisition: a study of young children learning new names. *Br. J. Psychol.* 8, 439–454. doi: 10.1111/j.2044-8295.1990.tb02371.x

Gathercole, S. E., Service, E., Hitch, G. J., Adams, A. M., and Martin, A. J. (1999). Phonological short-term memory and vocabulary development: further evidence on the nature of the relationship. *Appl. Cogn. Psychol.* 13, 65–77. doi: 10.1002/(SICI)1099-0720(199902)13:1<65::AID-ACP548>3.0.CO;2-O

Golestani, N., Molko, N., Dehaene, S., LeBihan, D., and Pallier, C. (2007). Brain structure predicts the learning of foreign speech sounds. *Cereb. Cortex* 17, 575–582. doi: 10.1093/cercor/bhk001

Golestani, N., and Pallier, C. (2007). Anatomical correlates of foreign speech sound production. *Cereb. Cortex* 17, 929–934. doi: 10.1093/cercor/bhl003

Golestani, N., Paus, T., and Zatorre, R. J. (2002). Anatomical correlates of learning novel speech sounds. *Neuron* 35, 997–1010. doi: 10.1016/S0896-6273(02)00862-0

Golestani, N., Price, C. J., and Scott, S. K. (2011). Born with an ear for dialects? Structural plasticity in the expert phonetician brain. *J. Neurosci.* 31, 4213–4220. doi: 10.1523/JNEUROSCI.3891-10.2011

Golestani, N., and Zatorre, R. J. (2009). Individual differences in the acquisition of second language phonology. *Brain Lang.* 109, 55–67. doi: 10.1016/j.bandl.2008.01.005

Gordon, E. E. (1980). The assessment of music aptitudes of very young children. *Gift. Child Q.* 24, 107–111. doi: 10.1177/001698628002400303

Gordon, E. E. (2001). *Music Aptitude and Related Tests*. Chicago, IL: GIA Publications, Inc.

Granena, G. (2013). "Cognitive aptitudes for second language learning and the LLAMA language aptitude test," in *Sensitive Periods, Language Aptitude, and Ultimate L2 Attainment*, eds G. Granena and M. Long (Amsterdam: John Benjamins Publishing), 105–130.

Granena, G., and Long, M. (2013). *Sensitive Periods, Language Aptitude, and Ultimate L2 Attainment*. Amsterdam: John Benjamins Publishing.

Hackett, T. A. (2009). "The evolution of the primate and human auditory system," in *Evolutionary Neuroscience*, ed. J. H. Haas (Oxford: Academic Press), 893–905.

Heschl, R. L. (1878). *Über Die Vordere Quere Schläfenwindung Des Menschlichen Grosshirns*. Wien: Braumüller.

Hu, X., Ackermann, H., Martin, J. A., Erb, M., Winkler, S., and Reiterer, S. M. (2013). Language aptitude for pronunciation in advanced second language (L2) learners: behavioural predictors and neural substrates. *Brain Lang.* 127, 366–376. doi: 10.1016/j.bandl.2012.11.006

Jacquemot, C., and Scott, S. K. (2006). What is the relationship between phonological short-term memory and speech processing? *Trends Cogn. Sci.* 10, 480–486. doi: 10.1016/j.tics.2006.09.002

Jäncke, L. (2012). The relationship between music and language. *Front. Psychol.* 3:123. doi: 10.3389/fpsyg.2012.00123

Jilka, M. (2009). "Talent and proficiency in language," in *Language Talent and Brain Activity*, eds G. Dogil and S. M. Reiterer (Berlin: Mouton de Gruyter), 1–16.

Kepinska, O. (2017). *The Neurobiology of Individual Differences in Grammar Learning*. Ph.D. dissertation, University of Leiden, Leiden.

Kepinska, O., de Rover, M., Caspers, J., and Schiller, N. O. (2017a). On neural correlates of individual differences in novel grammar learning: an fMRI study. *Neuropsychologia* 98, 156–168. doi: 10.1016/j.neuropsychologia.2016.06.014

Kepinska, O., de Rover, M., Caspers, J., and Schiller, N. O. (2017b). Whole-brain functional connectivity during acquisition of novel grammar: distinct functional networks depend on language learning abilities. *Behav. Brain Res.* 320, 333–346. doi: 10.1016/j.bbr.2016.12.015

Klein, E. C. (1995). Second versus third language acquisition: is there a difference? *Lang. Learn.* 45, 419–466. doi: 10.1111/j.1467-1770.1995.tb00448.x

Kocić, A. (2010). The effects of aptitude on language learning. *Komun. Kult.* 1, 234–243.

Koelsch, S. (2005). Neural substrates of processing syntax and semantics in music. *Curr. Opin. Neurobiol.* 15, 1–6. doi: 10.1016/j.conb.2005.03.005

Kormos, J., and Sáfár, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive

language learning. *Bilingualism* 11, 261–271. doi: 10.1017/S13667289080 03416

Kraus, N., and Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nat. Rev. Neurosci.* 11, 599–605. doi: 10.1038/nrn2882

Lee, L., and Lin, S. C. (2015). The impact of music activities on foreign language, English learning for young children. *J. Eur. Teach. Educ. Netw.* 10, 13–23.

Li, S. (2015). The associations between language aptitude and second language grammar acquisition: a meta-analytic review of five decades of research. *Appl. Linguist.* 36, 385–408. doi: 10.1093/applin/amu054

Li, S. (2016). The construct validity of language aptitude. *Stud. Second Lang. Acquis.* 38, 801–842. doi: 10.1017/S027226311500042X

Linck, J. A., Hughes, M. M., Campbell, S., Silbert, N. H., Tare, M., Jackson, S. R., et al. (2013). Hi-LAB: a new measure of aptitude for high-level language proficiency. *Lang. Learn.* 63, 530–566. doi: 10.1111/lang.12011

Linck, J. A., Osthus, P., Koeth, J. T., and Bunting, M. F. (2014). Working memory and second language comprehension and production: a meta-analysis. *Psychon. Bull. Rev.* 21, 861–883. doi: 10.3758/s13423-013-0565-2

Ludke, K. M. (2010). *Songs and Singing in Foreign Language Learning.* Ph.D. dissertation. University of Edinburgh, Edinburgh.

Ludke, K. M., Ferreira, F., and Overy, K. (2014). Singing can facilitate foreign language learning. *Mem. Cogn.* 42, 41–52. doi: 10.3758/s13421-013-0342-5

Magne, C., Schön, D., and Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: behavioral and electrophysiological approaches. *J. Cogn. Neurosci.* 18, 199–211. doi: 10.1162/jocn.2006.18.2.199

Marie, D., Maingault, S., Crivello, F., Mazoyer, B., and Tzourio-Mazoyer, N. (2016). Surface-based morphometry of cortical thickness and surface area associated with heschl's gyri duplications in 430 healthy volunteers. *Front. Hum. Neurosci.* 10:69. doi: 10.3389/fnhum.2016.00069

McDonald, J. H. (2014). *Handbook of Biological Statistics,* 3rd Edn. Baltimore, MD: Sparky House Publishing.

McGettigan, C., and Scott, S. K. (2012). Cortical asymmetries in speech perception: what's wrong, what's right and what's left? *Trends Cogn. Sci.* 16, 269–276. doi: 10.1016/j.tics.2012.04.006

Meara, P. (2005). *LLAMA Language Aptitude Tests: The Manual.* Swansea: Lognostics.

Milovanov, R. (2009). *The Connectivity of Musical Aptitude and Foreign Language Learning Skills: Neural and Behavioural Evidence.* Turku: University of Turku.

Milovanov, R., Huotilainen, M., Esquef, P. A. A., Välimäki, V., Alku, P., and Tervaniemi, M. (2009). The role of musical aptitude and language skills in preattentive duration determination in school-aged children. *Neurosci. Lett.* 460, 161–165. doi: 10.1016/j.neulet.2009.05.063

Milovanov, R., Huotilainen, M., Velimäki, V., Esquef, P. A. A., and Tervaniemi, M. (2008). Musical aptitude and second language pronunciation skills in school-aged children: neural and behavioral evidence. *Brain Res.* 1194, 81–89. doi: 10.1016/j.brainres.2007.11.042

Milovanov, R., and Tervaniemi, M. (2011). The interplay between musical and linguistic aptitudes: a review. *Front. Psychol.* 2:321. doi: 10.3389/fpsyg.2011. 00321

Milovanov, R., Tervaniemi, M., and Gustafsson, M. (2004). "The impact of musical aptitude in foreign language acquisition," in *Proceedings of the 8th International Conference on Music Perception & Cognition,* eds S. D. Lipscomb, R. Ashley, R. O. Gjerdingen, and P. Webster (Adelaide, SA: Causal Productions).

Milovanov, R., Pietilä, P., Tervaniemi, M., and Esquef, P. A. A. (2010). Foreign language pronunciation skills and musical aptitude: a study of Finnish adults with higher education. *Learn. Individ. Diff.* 20, 56–60. doi: 10.1016/j.lindif.2009. 11.003

Miyake, A., and Friedman, N. P. (1998). "Individual differences in second language proficiency: working memory as language aptitude," in *Foreign Language Learning - Psycholinguistic Studies on Training and Retention,* eds A. F. Healy and L. E. Bourne Jr. (Mahwah, NJ: Lawrence Erlbaum), 339–364.

Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., and Zilles, K. (2001). Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system. *Neuroimage* 13, 684–701. doi: 10.1006/nimg.2000.0715

Nardo, D., and Reiterer, S. M. (2009). "Musicality and phonetic language aptitude," in *Language Talent and Brain Activity,* eds G. Dogil and S. M. Reiterer (Berlin: Mouton de Gruyter), 213–256.

Patel, A. D. (2012). "Language, music, and the brain: a resource-sharing framework," in *Language and Music as Cognitive Systems,* eds P. Rebuschat, M. Rohrmeier, J. Hawkins, and I. Cross (Oxford: Oxford University Press), 204–223.

Patel, D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Front. Psychol.* 2:142. doi: 10.3389/fpsyg.2011. 00142

Penhune, V. B., Zatorre, R. J., MacDonald, J. D., and Evans, A. C. (1996). Interhemispheric anatomical differences in human primary auditory cortex: probabilistic mapping and volume measurement from magnetic resonance scans. *Cereb. Cortex* 6, 661–672. doi: 10.1093/cercor/6.5.661

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. *Speech Commun.* 41, 245–255. doi: 10.1016/S0167-6393(02)00107-3

Prat, C. S., Yamasaki, B. L., Kluender, R. A., and Stocco, A. (2016). Resting-state qEEG predicts rate of second language learning in adults. *Brain Lang.* 15, 44–50. doi: 10.1016/j.bandl.2016.04.007

Price, C. J. (2010). The anatomy of language: a review of 100 fMRI studies published in 2009. *Ann. N. Y. Acad. Sci.* 1191, 62–88. doi: 10.1111/j.1749-6632.2010. 05444.x

Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage* 62, 816–847. doi: 10.1016/j.neuroimage.2012.04.062

Purves, D., Augustine, G. J., Fitzpatrick, D., Katz, L. C., La Mantia, A.-S., McNamara, J. O., et al. (2001). *Neuroscience.* Sunderland, MA: Sinauer Associates.

Rademacher, J., Galaburda, A. M., Kennedy, D. N., Filipek, P. A., and Caviness, V. S. J. (1992). Human cerebral cortex: localization, parcellation, and morphometry with magnetic resonance imaging. *J. Cogn. Neurosci.* 4, 352–374. doi: 10.1162/jocn.1992.4.4.352

Rademacher, J., Morosan, P., Schormann, T., Schleicher, A., Werner, C., Freund, H.-J., et al. (2001). Probabilistic mapping and volume measurement of human primary auditory cortex. *Neuroimage* 13, 669–683. doi: 10.1006/nimg.2000. 0714

Reiterer, S. M., Hu, X., Erb, M., Rota, G., Nardo, D., Grodd, W., et al. (2011). Individual differences in audio-vocal speech imitation aptitude in late bilinguals: functional neuro-imaging and brain morphology. *Front. Psychol.* 2:271. doi: 10.3389/fpsyg.2011.00271

Robinson, P. (2001). Individual differences, cognitive abilities, aptitude complexes and learning conditions in second language acquisition. *Second Lang. Res.* 17, 368–392. doi: 10.1177/026765830101700405

Robinson, P. (2002). "Effects of individual differences in intelligence, aptitude, and working memory on adult incidental SLA: a replication and extension of Reber, Walkenfield and Hernstad (1991)," in *Individual Differences and Instructed Language Learning,* ed. P. Robinson (Amsterdam: John Benjamins Publishing), 211–265.

Robinson, P. (2005). Aptitude and second language acquisition. *Annu. Rev. Appl. Linguist.* 25, 46–73. doi: 10.1017/S0267190505000036

Robinson, P. (2012). "Individual differences, aptitude complexes, SLA processes, and aptitude test development," in *New Perspectives on Individual Differences in Language Learning and Teaching,* ed. M. Pawlak (Berlin: Springer), 57–75.

Rogers, V. E., Meara, P., Aspinall, R., Fallon, L., Goss, T., Keey, E., et al. (2016). Testing aptitude. *EUROSLA Yearb.* 16, 179–210. doi: 10.1075/eurosla.16.07rog

Rota, G., and Reiterer, S. M. (2009). "Cognitive aspects of pronunciation talent: how empathy, mental flexibility, working memory and intelligence interact with phonetic talent," in *Language Talent and Brain Activity,* eds G. Dogil and S. M. Reiterer (Berlin: Mouton de Gruyter), 67–96.

Sáfár, A., and Kormos, J. (2008). Revisiting problems with foreign language aptitude. *IRAL* 46, 113–136. doi: 10.1515/IRAL.2008.005

Sawyer, M., and Ranta, L. (2001). "Aptitude, individual differences, and instructional design," in *Cognition and Second Language Instruction,* ed. P. Robinson (Cambridge: Cambridge University Press), 319–353. doi: 10.1017/ CBO9781139524780.013

Schneider, P., Andermann, M., Wengenroth, M., Goebel, R., Flor, H., Rupp, A., et al. (2009). Reduced volume of Heschl's Gyrus in tinnitus. *Neuroimage* 45, 927–939. doi: 10.1016/j.neuroimage.2008.12.045

Schneider, P., Scherg, M., Dosch, H. G., Specht, H. J., Gutschalk, A., and Rupp, A. (2002). Morphology of Heschl's gyrus reflects enhanced activation

in the auditory cortex of musicians. *Nat. Neurosci.* 5, 688–694. doi: 10.1038/nn871

Schneider, P., Sluming, V., Roberts, N., Bleeck, S., and Rupp, A. (2005). Structural, functional, and perceptual differences in Heschl's Gyrus and musical instrument preference. *Ann. N. Y. Acad. Sci.* 1060, 387–394. doi: 10.1196/annals.1360.033

Schön, D., Magne, C., and Besson, M. (2004). The music of speech: music training facilitates pitch processing in both music and language. *Psychophysiology* 41, 341–349. doi: 10.1111/1469-8986.00172.x

Seither-Preisler, A., Parncutt, R., and Schneider, P. (2014). Size and synchronization of auditory cortex promotes musical, literacy, and attentional skills in children. *J. Neurosci.* 34, 10937–10949. doi: 10.1523/JNEUROSCI.5315-13.2014

Serrallach, B., Groß, C., Bernhofs, V., Engelmann, D., Benner, J., Gündert, N., et al. (2016). Neural biomarkers for dyslexia, ADHD, and ADD in the auditory cortex of children. *Front. Neurosci.* 10:324. doi: 10.3389/fnins.2016.00324

Shabani, M. B., and Torkeh, M. (2014). The relationship between musical intelligence and foreign language learning: the case of Iranian learners of English. *Int. J. Appl. Linguist. Engl. Lit.* 3, 26–32. doi: 10.7575/aiac.ijalel.v.3n.3p.26

Singleton, D. (2017). Language aptitude: desirable trait or acquirable attribute? *Stud. Second Lang. Learn. Teach.* 7, 89–103. doi: 10.14746/ssllt.2017.7.1.5

Skehan, P. (1986). "Where does language aptitude come from?," in *Spoken Language*, ed. P. M. Meara (London: Centre for Information and Language Teaching), 95–113.

Skehan, P. (2002). "Theorising and updating aptitude," in *Individual Differences and Instructed Language Learning*, ed. P. Robinson (Amsterdam: John Benjamins), 69–95.

Slevc, L. R., and Miyake, A. (2006). Individual differences in second-language proficiency. *Psychol. Sci.* 17, 675–681. doi: 10.1111/j.1467-9280.2006.01765.x

Sparks, R., and Ganschow, L. (2001). Aptitude for learning a foreign language. *Annu. Rev. Appl. Linguist.* 21, 90–111. doi: 10.1017/S026719050100006X

Sparks, R. L., Humbach, N., Patton, J. O. N., and Ganschow, L. (2011). Subcomponents of second-language aptitude and second-language proficiency. *Modern Lang. J.* 95, 253–273. doi: 10.1111/j.1540-4781.2011.01176.x

Spolsky, B. (1995). Prognostication and language aptitude testing, 1925-62. *Lang. Test.* 12, 321–340. doi: 10.1177/026553229501200304

Stansfield, C. W., and Reed, D. J. (2004). The story behind the modern language aptitude test: an interview with John B. Carroll (1916–2003). *Lang. Assess. Q.* 1, 43–56. doi: 10.1207/s15434311laq0101_4

Stern, E., and Neubauer, A. (2013). *Lernen Macht Intelligent: Warum Begabung Gefördert Werden Muss.* München: DVA.

Talairach, J., and Tournoux, P. (1988). *Co-planar Stereotaxic Atlas of the Human Brain. 3-Dimensional Proportional System: An Approach to Cerebral Imaging.* New York, NY: Thieme.

Thompson, A. S. (2013). The interface of language aptitude and multilingualism: reconsidering the bilingual/multilingual dichotomy. *Modern Lang. J.* 97, 685–701. doi: 10.1111/j.1540-4781.2013.12034.x

Van den Noort, M. W., Bosch, P., and Hugdahl, K. (2006). Foreign language proficiency and working memory capacity. *Eur. Psychol.* 11, 289–296. doi: 10.1027/1016-9040.11.4.289

Vangehuchten, L., Verhoeven, V., and Thys, P. (2015). Pronunciation proficiency and musical aptitude in Spanish as a foreign language: results of an experimental research project. *Rev. Lingüíst. Lenguas Aplicadas* 10, 90–100. doi: 10.4995/rlyla.2015.3372

Warrier, C., Wong, P., Penhune, V., Zatorre, R., Parrish, T., and Kraus, N. (2012). Relating structure to function: Heschl's Gyrus and acoustic processing. *J. Neurosci.* 29, 61–69. doi: 10.1523/JNEUROSCI.3489-08.2009

Wen, Z. (2012). Foreign language aptitude. *ELT J.* 66, 233–235. doi: 10.1093/elt/ccr068

Wen, Z. (2016). *Working Memory and Second Language Learning: Towards an Integrated Approach.* Bristol: Multilingual Matters.

Wen, Z., and Skehan, P. (2011). A new perspective on foreign language aptitude research: building and supporting a case for "working memory as language aptitude". *Ilha Do Desterro J. Engl. Lang. Lit. Engl. Cult. Stud.* 60, 15–43. doi: 10.5007/2175-8026.2011n60p015

Wen, Z., Skehan, P., and Biedroń, A. (2017). Foreign language aptitude theory: yesterday, today and tomorrow. *Lang. Teach.* 50, 1–31. doi: 10.1017/S0261444816000276

Wengenroth, M., Blatow, M., Bendszus, M., and Schneider, P. (2010). Leftward lateralization of auditory cortex underlies holistic sound perception in Williams syndrome. *PLOS ONE* 5:e12326. doi: 10.1371/journal.pone.0012326

Wengenroth, M., Blatow, M., Heinecke, A., Reinhardt, J., Stippich, C., Hofmann, E., et al. (2013). Increased volume and function of right auditory cortex as a marker for absolute pitch. *Cereb. Cortex* 24, 1127–1137. doi: 10.1093/cercor/bhs391

Winke, P. (2013). An investigation into second language aptitude for advanced Chinese language learning. *Modern Lang. J.* 97, 109–130. doi: 10.1111/j.1540-4781.2013.01428.x

Wong, P., Perrachione, T. K., and Parrish, T. B. (2007). Neural characteristics of successful and less successful speech and word learning in adults. *Hum. Brain Mapp.* 28, 995–1006. doi: 10.1002/hbm.20330

Wong, P. C. M., Warrier, C. M., Penhune, V. B., Roy, A. K., Sadehh, A., Parrish, T. B., et al. (2008). Volume of left Heschl's Gyrus and linguistic pitch learning. *Cereb. Cortex* 18, 828–836. doi: 10.1093/cercor/bhm115

Xiang, H. (2012). *The Language Networks of the Brain*. Ph.D. dissertation, Radboud Universiteit Nijmegen, Nijmegen.

Xiang, H., Dediu, D., Roberts, L., Oort, E. V., Norris, D. G., and Hagoort, P. (2012). The structural connectivity underpinning language aptitude, working memory, and IQ in the perisylvian language network. *Lang. Learn.* 62, 110–130. doi: 10.1111/j.1467-9922.2012.00708.x

Zatorre, R. J., Belin, P., and Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends Cogn. Sci.* 6, 37–46. doi: 10.1016/S1364-6613(00)01816-7

Predictors of Successful Learning in Multilingual Older Adults Acquiring a Majority Language

Henrike K. Blumenfeld[1]*, Sim J. R. Quinzon[1], Cindy Alsol[1,2] and Stephanie A. Riera[1]

[1] School of Speech, Language and Hearing Sciences, San Diego State University, San Diego, CA, United States, [2] Somali Family Service of San Diego, San Diego, CA, United States

Understanding language learning in later life can elucidate how linguistic experiences and age-specific cognitive skills can be leveraged for language acquisition, providing insight into how lifelong experiences configure our learning capacity. In this study, we examined to what extent acquisition and maintenance of a non-native language (English) is scaffolded by cognitive skills and previous linguistic experiences in older adults; and to what extent these cognitive/linguistic factors predict older learners' success in acquiring novel functional language. We recruited 53 participants who were native speakers of Mandarin, Spanish, Tagalog, and Somali, had continued to learn English as adults, and were currently exposed to majority-English contexts. To identify contributors to participants' English skills, we administered a language history and self-reported proficiency interview, brief cognitive testing, and verbal fluency tasks in L1 and English. We found that digit span and orientation measures were cognitive predictors of English proficiency, while similarity of known languages to English, L1 skills, and English language exposure were linguistic predictors of English skills. To examine participants' ability to maintain language knowledge and to learn new functional English, we also conducted a preliminary longitudinal service-based study in a subset of 19 participants using our *Specific-Purpose English Communication System for Seniors* (*SPECSS*) curriculum. In this subset of SPECSS learners, we identified digit span and orientation, but not age, as cognitive predictors of short-term language maintenance. Further, better novel English learning as a result of our curriculum was observed in learners whose other known languages were *less* similar to English. Findings inform best practices in developing language curricula for older adults, and help generate new hypotheses on preparedness for language learning across the adult lifespan with a possible interaction between cognitive skills and transfer of knowledge from previous languages in multilingual older learners.

Keywords: adult language learning, cognitive aging, multilingualism, language transfer, cross-linguistic influence, language experience

INTRODUCTION

Increased age has long been thought of as limiting individuals' abilities to learn new languages, consistent with age-related changes in memory (e.g., Ullman, 2001; Janacsek et al., 2012) as well as in neural plasticity (e.g., Lillard and Erisir, 2011). Yet, cognitive benefits (Bak et al., 2016), neural reorganization (Mohr et al., 2014), and learning success (Marcotte and Ansaldo, 2014)

have been demonstrated with language training in older adults. The literature on older adults' language learning remains sparse (Blumenfeld, 2012; Antoniou et al., 2013; Marcotte and Ansaldo, 2014; Bak et al., 2016) and clear practical and theoretical needs exist for a better understanding of language learning capacity across the lifespan. For example, of the US population above age 60, 15% speak a language other than English at home and, of these individuals, 58% speak English "less than very well" (Ryan, 2013). Low proficiency in the majority language has been shown to reduce health outcomes (Ponce et al., 2006; Mui et al., 2007) and well-being (Ding and Hargraves, 2009), creating a need for age/cognition-appropriate language curricula that can enhance the functional English skills of older adults. Further, understanding language learning in later life can elucidate how extensive linguistic experiences and age-specific cognitive skills can be leveraged for novel language acquisition, thus providing insight into how lifelong experiences configure our language learning capacity. Here, we examine cognitive and linguistic predictors of language attainment in a diverse group of older adults who are late learners of English, and report on an initial examination of language maintenance and novel learning in a subset of this group that can inform benefits of and approaches to language learning in older adults.

Hurdles to adult foreign language learning include greater entrenchment of already-acquired linguistic knowledge, potentially making it more challenging for adult learners to re-structure representations during novel learning. Relatedly, robust previous language representations may result in negative transfer of previous knowledge to the new language, yielding errors, and incomplete acquisition. Moreover, lack of social opportunities to use the new language may limit the extent of immersion. On the flip side, factors that may optimize adult language acquisition include positive transfer of previously acquired knowledge to the new language and internalization of novel linguistic information through continued immersion (Unified Model of Second Language Acquisition, MacWhinney, 2005, 2012). These hurdles and protective factors provide a framework to examine cognitive and experiential predictors of language attainment and maintenance in older adults, with a focus on three interrelated factors: age-related cognitive processes, previous knowledge of L1 and other languages, and experience in the new language.

Language learning involves a number of cognitive skills, including the ability to hold novel sound representations in phonological short-term memory (e.g., Papagno et al., 1991; Papagno and Vallar, 1995; Kaushanskaya, 2012) and working memory for later integration (e.g., Miyake and Friedman, 1998) and consolidation (Whitfield and Goberman, 2017). In addition, adult foreign language learning has been shown to involve integration of novel and previous knowledge through associations between translation equivalents (e.g., Kroll and Stewart, 1994), blending of semantic content (De Groot, 1992), and both positive and negative transfer between overlapping and distinct aspects of the previous and novel languages (MacWhinney, 2012). Therefore, language learners must walk a fine line between allowing co-activation of their languages for integration and transfer, and inhibiting previous languages to allow novel learning and processing. Consistently, evidence from a number of studies with

younger adults suggests that cognitive resources are recruited to manage interference from non-target languages in individuals who are learning a novel language (e.g., Raboyeau et al., 2010; Bartolotti et al., 2011).

With cognitive aging, declines have been identified in processes that underlie language learning, including phonological short-term memory and working memory (e.g., Gregoire and Van der Linden, 1997), encoding (e.g., Craik, 2002), and consolidation of new memories (e.g., Meyer and Federmeier, 2010), as well as inhibitory control (e.g., Lustig et al., 2007). Consistently, in the linguistic domain, older adults have been shown to be less likely than younger adults to recruit cognitive processes for competition resolution (e.g., Blumenfeld et al., 2016b), benefit more from the presence of a semantic context during ambiguity resolution (Lee and Federmeier, 2011), and are less likely to re-interpret linguistic information (e.g., Meyer and Federmeier, 2010). These age-related changes allow for a series of predictions on language learning in older adults, including less efficient learning because of decline in executive function, and a potential shift to alternative cognitive pathways. Indeed, Marcotte and Ansaldo (2014) found that older adults do succeed at language learning but do so with more practice and through different learning strategies. Marcotte and Ansaldo (2014) found similar learning outcomes when younger and older French-speakers learned Spanish. However, the older adults required more time (25 days instead of 14 days in younger adults) to reach ceiling in learning Spanish words. In addition, Marcotte and Ansaldo's neuroimaging findings revealed that older learners relied more on episodic memory and visual learning pathways than their younger peers who relied more on frontal cognitive control networks. In fact, in a recent review, Amer et al. (2016) have argued that older adults' greater reliance on previously encoded information in new contexts relates directly to their reduced reliance on cognitive control. Older learners may thus show longer learning trajectories with increased reliance on previous knowledge.

Even younger adult learners have been shown to rely heavily on previous linguistic knowledge when acquiring a novel language. A robust research base exists on language transfer as a significant contributor to language learning [e.g., Lotto and De Groot, 1998; Sparks et al., 2009; Morett and MacWhinney, 2013; Antoniou et al., 2014; Bartolotti and Marian, 2016; for a recent review see Hirosh and Degani (2017)]. For example, Antoniou et al. (2014) found that learners who knew Mandarin attained better learning outcomes for an artificial language that contained a retroflex contrast found in Mandarin; similarly, learners who knew Korean outperformed others in learning a language with a lenition contrast found in Korean. Consistently, Bartolotti and Marian (2016) taught fluent speakers of English and German a novel artificial language that had overlap with both their L1 and L2, and found that both previously learned languages contributed to success with the novel language. Finally, in proficient speakers, structurally similar aspects of languages continue to provide cross-linguistic scaffolding for processing in both bilingual contexts (e.g., Costa et al., 2005; Schoonbaert et al., 2007; Blumenfeld et al., 2016a; Potapova et al., 2016) and multilingual contexts (e.g., Lemhöfer et al., 2004). It has thus

been well established in young adults that previously known languages provide an experiential baseline that can facilitate the acquisition of novel languages, both through direct transfer of knowledge and through potential honing of cognitive skills that underlie learning (e.g., Hirosh and Degani, 2017).

Consistently with predictions that older adults may rely heavily on previous knowledge, Marcotte and Ansaldo (2014) found in their word-level training study, teaching Spanish words to younger and older French monolinguals, that older learners had more robust cognate effects than the younger learners. This effect was driven by particular challenges in the initial learning of non-cognate words in the older learners, and was no longer significant once learners had reached ceiling. These findings suggest that longer learning phases in older adults are particularly present when novel L2 targets must be mastered that do not resemble previous knowledge. Therefore, current research is consistent with the expectation that older learners may be particularly reliant on transfer of previous language knowledge during L2 acquisition.

Nevertheless, findings from Siyambalapitiya et al. (2009), and from an older adult control group in Roberts and Deslauriers (1999) suggest that bilingual older adults may not show consistent cognate processing advantages, perhaps because of the cognitive costs associated with co-activation of two languages (Hughes and Tainturier, 2015). Despite Marcotte and Ansaldo (2014) findings, it is conceivable that, with reduced cognitive control skills (e.g., Lustig et al., 2007), older adults may at times struggle in acquiring linguistic information that is somewhat similar to previous knowledge (thus encouraging co-activation with previously known languages) yet has different features (thus requiring cognitive muting of previously known languages). Therefore, additional research is warranted into the nature of language transfer during learning in older adults to further delineate cognitive and experiential contributing factors.

In addition to positive transfer, another protective factor for adult learning success identified within the Unified Second Language Acquisition model is continued immersion in the new language (MacWhinney, 2012). The importance of continued language use with age is apparent in the literature from monolinguals. For example, Barresi et al. (1998) found in a longitudinal study that older individuals who reported living in a household with other adults showed better naming performance while those who reported high-passive language exposure through television showed lower performance. In younger bilingual adults and language learners, language exposure has similarly emerged as an important predictor of abilities (e.g., Marian et al., 2007; Linck et al., 2009) and may play an important role in the maintenance of L2 in older adults (e.g., Nanchen et al., 2017). It has been suggested that continued use of a language provides continued activation and strengthens its representations, creating language-specific *resonance* that boosts the network underlying the novel language and reduces interference from other languages (MacWhinney, 2012). It is thus likely that, with slower encoding and learning, and with fewer cognitive resources available to mitigate interference from more proficient languages, continued immersion is particularly critical for older learners.

With more effortful learning, strategies for foreign language acquisition have been shown to shift in older learners. Older learners have been found to recruit more cortical regions underlying visual imagery and episodic memories compared with younger peers (lingual gyrus, precuneus, cuneus, Marcotte and Ansaldo, 2014), a finding that was interpreted as a stronger reliance on visual semantic information provided during learning (Stuart et al., 2006), with less reliance on cognitive control circuitry. Indeed, semantic memory has been found to be especially well-preserved with age (e.g., Reuter-Lorenz et al., 2000), and previously established semantic processes may thus serve as scaffolding for learning of novel information in older adults. Therefore, both learning speed and pathways are likely to differ across younger and older learners, and classrooms that are age-specific may be most appropriate to fully accommodate older learners (Marinova-Todd et al., 2000). Given these findings of language learning mechanisms, learning materials where familiar semantic contexts of use are clearly established and visually presented may be especially beneficial for older learners. Thematically organized practical content is also likely to be more immediately useful to learners (e.g., Antoniou et al., 2013) and may thus be especially critical for older learners who acquire language more effortfully.

In the present study, we examined whether previously established cognitive and linguistic factors that contribute to language learning would jointly contribute to the ability to gain language skills in a multilingual group of older adults. We were particularly interested in whether the nature of previous language learning would influence mastery of English in this group of non-native speakers and whether such previous linguistic experience would influence short-term language maintenance and continued guided learning of functional English through a multi-week tailored curriculum we designed, our *Specific-Purpose English Communication System for Seniors* (*SPECSS*). For purposes of this study, we operationally defined short-term language maintenance as the retention of language knowledge as measured before and after participation in the SPECSS curriculum.

We recruited a group of older adults who were native speakers of Mandarin, Spanish, Tagalog, and Somali, who continued to learn English as adults and were currently exposed to majority-English contexts in the USA. We hoped to identify contributors to participants' current English skills through a language history and self-reported proficiency interview, as well as through brief cognitive testing. Participants' language attainment was indexed through self-reports and through verbal fluency tasks in their L1 and in English. In addition, we conducted a longitudinal service-based study in a subset of these participants where we examined their ability to maintain and learn a functional English language curriculum. The English curriculum was tailored to the expected learning needs of older adults acquiring a majority language, and included six topic modules on communication basics, small talk, interacting with healthcare providers, emergencies, navigating the community, and grocery shopping. In addition, the teaching materials and approach were developed to accommodate expected learning styles of older adults, including increased opportunity for rehearsal of material, as well as easy access to native-language translation equivalents, a strategy that has

been shown to facilitate adult second language acquisition (e.g., Lotto and De Groot, 1998). The thematic organization of the curriculum and playing out of specific everyday situations was based on findings that retrieval from memory is easiest when the language and context at retrieval match those at encoding (e.g., Marian and Kaushanskaya, 2011). Therefore, the curriculum was designed to simulate real-life situations older adults might encounter, with functional target words and phrases to facilitate communication. Further, salient visual referents were provided in the materials given older adults' identified focus on perceptual information during learning (Stuart et al., 2006; Marcotte and Ansaldo, 2014).

We asked (1) to what extent acquisition of a low-proficiency non-native language (English) would be scaffolded by cognitive skills and previous linguistic experiences in older adults; and (2) to what extent these cognitive and linguistic factors would predict older learners' short-term language maintenance and success in acquiring novel functional language skills through a focused curriculum. We predicted that phonological short-term and working memory and attention, as well as amount of English exposure, would emerge as predictors of performance in English. Second, we hypothesized that the multilingual language learners whose previously known languages are similar to English might show the greatest English attainment, maintenance, and novel learning, because they can rely on language transfer. As an alternative prediction, Hirosh and Degani (2017) have recently argued that multilinguals with *less similar* previously known languages may have a novel language learning advantage because they are more likely to globally inhibit their previous unrelated languages. We expected that the initial language maintenance and learning data from participants who completed our SPECSS curriculum would provide insight on these alternative hypotheses to help formulate effective language curricula for older adults and to guide future research.

MATERIALS AND METHODS

Participants

Fifty-three older adult non-native speakers of English participated in this study (mean age = 72.92, SD = 6.72, range: 58–81 years; 34 female). This study was carried out in accordance with the recommendations of San Diego State University's Institutional Review Board. The protocol was approved by the San Diego State University Internal Review Board. Written informed consent was obtained from all participants. Participants were recruited at one of two local community centers and gave written informed consent in accordance with the Declaration of Helsinki. All participants spoke a native language other than English, and had no history of stroke.[1] The native languages spoken by the participants were Mandarin ($n = 19$), Spanish ($n = 12$), Somali ($n = 10$), and Tagalog ($n = 12$). Participants

---

[1]Ten participants reported having had a head injury in their adult life as a result of a fall ($n = 4$), car accident ($n = 4$), laboratory explosion ($n = 1$), or non-stated reason ($n = 1$), with four participants reporting loss of consciousness subsequent to injury. Since outcomes of all analyses remained the same when these 10 participants were omitted, we included all participants in the current cohort.

had an average of 11.27 years of formal education (SD = 5.94, range: 0–19 years) and had first been exposed to English at an average age of 32.88 (SD = 22.30, range: 5–74 years). Participants who reported exposure to English at or before age 7 ($n = 7$) reported ages of immigration to the USA well after childhood (mean age of immigration = 51.8 years, range: 20–72). These participants reported other languages as L1/home languages. While these participants reported being exposed to English in school, this English was limited (e.g., Bautista and Bolton, 2008). To obtain information on the language history and current language knowledge of participants, the Language Experience and Proficiency Questionnaire (LEAP-Q; Marian et al., 2007) was administered. To assess vocabulary in the native language and English, two semantic verbal fluency tasks (animals and groceries) were administered. All participants showed higher proficiency in their native language on the LEAP-Q and verbal fluency tasks (all $ps < 0.001$). See **Table 1** for a summary of participants' linguistic and cognitive profiles, and **Table 2** for a summary of languages spoken by the participants.

Of the 53 participants, 19 (12 female) were enrolled in classes using our *SPECSS* curriculum. These participants had a mean age of 67.74 (SD = 6.51, range: 58–81) and were native speakers of Mandarin ($n = 3$), Spanish ($n = 4$), Somali ($n = 10$), and Tagalog ($n = 2$). Participants had an average of 9.00 years of formal education (SD = 6.10, range: 0–18 years) and had an average age of 40.84 of first exposure to English (SD = 18.74, range: 7–63 years). Relative to the reference group, the learners reported later ages of first exposure to English, showed lower verbal fluency in English (animals and groceries), self-reported higher L1 proficiency, and showed lower L1 verbal fluency in the groceries category (all $ps < 0.05$). See **Table 1** for learner characteristics relative to

---

TABLE 1 | Participant characteristics of the reference group ($n = 53$) and the learner subset ($n = 19$).

	Reference group		Learners	
	Mean	**SD**	**Mean**	**SD**
Age of first exposure to English*	32.88	22.30	40.84	18.74
Current exposure to English	18.98	19.49	12.95	14.95
MoCA subtests				
Digit span (out of 2)	1.42	0.69	1.26	0.81
Orientation (out of 6)	5.68	0.67	5.47	0.96
Delayed memory recall (out of 5)	2.87	1.79	3.32	1.49
Proportion correct animal naming	0.79	0.28	0.86	0.26
English language skills				
Self-reported speaking (out of 10)	4.29	2.63	3.89	1.94
Self-reported comprehension (out of 10)	4.39	2.91	3.74	1.97
Self-reported reading (out of 10)	3.82	3.47	2.74	2.62
Verbal fluency—animals*	7.36	6.21	4.79	6.67
Verbal fluency—groceries*	8.91	5.88	5.79	6.74
L1 language skills				
Self-reported speaking (out of 10)*	8.63	1.44	9.32	1.06
Self-reported comprehension (out of 10)*	8.81	1.22	9.42	0.90
Self-reported reading (out of 10)	7.39	2.96	7.05	3.37
Verbal fluency—animals	14.72	4.36	15.11	4.21
Verbal fluency—groceries*	12.83	7.19	8.58	8.11

*Significant differences were observed between the reference and learner groups ($ps < 0.05$).

TABLE 2 | Languages spoken by participants and number of speakers, grouped by similarity to English on a 1 (least overlap) to 5 (most overlap) scale.

	L1 Mandarin	L1 Tagalog	L1 Spanish	L1 Somali	Total number of speakers
Outside of Indo-European Language Family with few borrowings (English-similarity score = 1)					
Mandarin	20	0	0	0	20
Arabic	0	1	0	6	7
Other Chinese Dialects	4	0	0	0	4
Amharic	0	0	0	1	1
Micronesian	0	1	0	0	1
Shanghainese	1	0	0	0	1
Tian-Jing Dialect	1	0	0	0	1
Outside of Indo-European Language Family with some borrowings (English-similarity score = 2)[a]					
Somali	0	0	0	10	10
Japanese	1	0	0	0	1
Kinamigin	0	1	0	0	1
Swahili	0	0	0	1	1
Outside of Indo-European Language Family with substantial borrowings (English-similarity score = 3)[b]					
Tagalog	0	9	0	0	9
Visayan languages (*Cebuano*, *Ilonggo*)	0	8	0	0	8
Ilocano	0	3	0	0	3
Bikol	0	1	0	0	1
Pangasinan	0	1	0	0	1
Indo-European Languages outside of the Germanic or Romance languages (English-similarity score = 4)					
Russian	14	0	0	0	14
Greek	0	0	1	0	1
Slovak	0	0	1	0	1
Within the Germanic or Romance Languages (English-similarity score = 5)					
Spanish	0	0	14	0	14
Italian	0	0	2	1	3
Chavacano[c]	0	1	0	0	1
French	0	0	1	0	1
German	0	0	1	0	1

[a]*Non-Indo-European languages categorized as having limited borrowings include Somali, with borrowings from English and Italian linked to European colonization (Somali, 2017); Japanese, with an estimated 10% of the lexicon borrowed from English (McKenzie, 2010, p. 14); Kinamigin, with documented Spanish presence in the Camiguin Island in the Bisayas region of the Philippines (Barreveld, 2001, p. 78); and Swahili, with borrowings from English where "contact with western civilization" existed, including in transportation, medicine, sports, and schools (Gower, 1952).*

[b]*Most of the major languages of the Philippines were categorized as having substantial borrowings, due to heavy lexical influence of Spanish (Lipski and Mühlhäusler, 1996; Rubino, 1997; Stolz, 2006; Mattes, 2014). The Spanish Colonial Era in the Philippines lasted from 1521 to 1898. These major Philippine languages also exhibit substantial borrowings from English (Rubino, 2001; Bernardo, 2004).*

[c]*Also known as Philippine Creole Spanish, Chavacano is the only Spanish-based creole in Asia (e.g., Lipski, 2013).*

the larger reference group. Overall, between-group comparisons point to cognitive and background similarities between the reference and learner groups, with lower English proficiency in the learner group.

Materials

All 53 participants were administered the LEAP-Q and verbal fluency tasks. In addition, cognitive skills were approximated using subtests of the *Montreal Cognitive Assessment* (MoCA; Nasreddine et al., 2005). In the 19 learners, knowledge of the SPECSS curriculum was also tested before and after they participated in classes.

Language Experience and Proficiency Questionnaire

Detailed information on the language history and proficiency of our older adult participants was obtained using the *LEAP-Q* (Marian et al., 2007). For all participants, a trained research

assistant who spoke the participant's native language gathered information during a 15- to 20-min structured oral interview that was closely based on the LEAP-Q. Participants provided basic information, such as age, education level, and exposure to English. Information specific to each of the participants' known languages, such as age of acquisition and self-ratings of language proficiency, were also obtained.

Verbal Fluency

Participants completed two verbal fluency tasks, including animal and grocery categories. Verbal fluency performance based on semantic category cues has been shown to index language proficiency in bilinguals (e.g., Gollan et al., 2002; Blumenfeld et al., 2016a). Animal and grocery categories were chosen since animals are a commonly used verbal fluency cue (e.g., Rosselli et al., 2000; Portocarrero et al., 2007; Bialystok et al., 2008) and the grocery cue was used to index participants' everyday language

use, where participants were instructed to list anything they could buy at the grocery store (e.g., Clark et al., 2009). Participants were verbally instructed to name as many items within each category as they could within 60 s without repetitions. Native-language versions of the verbal fluency tasks were administered as part of a testing session in participants' native language. English equivalents of the tasks were administered during a separate English session.

Cognitive Tasks

The *MoCA* (Nasreddine et al., 2005) was administered to gauge participants' cognitive performance. The MoCA is a well-established cognitive screening tool for older adults that covers a number of cognitive domains (executive function, memory, and language). Participants who spoke Mandarin, Spanish, or Tagalog as their native language completed the MoCA in their respective native languages. Participants in the Somali cohort completed a new Somali translation of the MoCA Basic (MoCA-B) that was deemed to be culturally acceptable, and appropriate given participants' lower education levels (Julayanont et al., 2015). Participants who were administered the MoCA-B were also given the forward and backward digit span subtest from the MoCA. Forward digit span is a measure of phonological short-term memory (i.e., the ability to retain and rehearse auditory stimuli), while backward digit span indexes working memory (e.g., Julayanont et al., 2012), and both measures have been found to underlie language learning (Papagno et al., 1991; Miyake and Friedman, 1998). Only subtests that had been completed by all participants were included in analyses, including attention (forward and backward digit span), orientation to time, date, and place, memory (delayed recall), and naming. The orientation subtest provides a measure of participants' awareness of where they are and what time and date it is, and such questions are typically included in cognitive assessment of older adults to index daily functioning (Julayanont et al., 2012); the delayed memory recall measure indexed participants' ability to encode words and retrieve them after a short time interval, and naming indexed knowledge and retrieval of core vocabulary. Instructions were given in the native language except for one participant in the reference group who preferred to take the test in English. Given participants' wide range in educational attainment, reported reading skills, and experiences with formal academically based tests, and given that the Somali version of the MoCA-B was a novel translation without validation data, scores of MoCA subtests were only used to approximate individual differences in cognitive skills across the participant group.

SPECSS Curriculum

All 19 participants in the learning component of the study received a $9 \times 7 \times 1.5$ inch portable ring binder containing the full curriculum to serve as a consistent memory and visual aid. Targets to be learned in English were presented with images and large-font text on one side of each page, and corresponding images and translations in participants' respective native language were on the flip side of that page. The binder was organized into topic modules deemed useful to participants based on feedback from staff at two senior centers (including social, nutrition, and community health workers). For a sample page from the binder, see **Figure 1**.

Topics were categorized into six modules including *Basics* (numbers, time, months/days of the week, directions, pronouns, and greetings), *Small Talk* (feelings, services, activities, and scheduling appointments), *Interacting with Healthcare Providers* (common patient history questions/answers, health conditions, professionals, medications, symptoms, allergies, body parts, and devices), *Emergencies* (types of emergencies such as medical, fire, etc.; calling for help and alerting others to emergencies; answering questions about what happened), *In the Community* (post office, transportation, requesting a translator, asking for directions, and phone etiquette), and *Groceries and Shopping* (grocery items; asking for help, price, and available discounts; and payment). The curriculum contained 69 pages, covering a total of 412 vocabulary items and phrases.

Translations of the curriculum into Spanish, Mandarin, Tagalog, and Somali were conducted through forward and backward translation procedures and checking of the materials by multiple proficient speakers of each language. Data from the Somali-speaking cohort were collected after data from the other participants had already been accrued, and minor modifications were made to materials to ensure cultural congruence for the Somali cohort: cartoon images of emotions were replaced with photos of a real person acting out the emotions due to the lack of familiarity with cartoon images in this group. All content remained the same across cohorts.

Procedures

After participants gave informed consent, assessments were administered individually by trained bilingual researchers in quiet testing rooms at the two local community centers where participants had been recruited. Native-language and English tasks were administered in separate sessions, with native-language sessions conducted first since this was participants' more dominant language. Participants were offered participation in the language learning component of the study, and the 19 seniors who agreed to enroll returned for individual baseline sessions where their knowledge of the curriculum was evaluated. During these baseline sessions, participants were shown the native-language sides of the curriculum pages from the SPECSS binder and were asked to translate target items to English equivalents. For participants requiring assistance, researchers read the target items for them in the native language. Following the baseline sessions, learners were given their personal SPECSS binder and enrolled in the SPECSS English classes, which were taught by trained bilingual researchers with teaching experience across a duration ranging from 12 (Somali cohort) to 21 weeks (Mandarin cohort). Learners participated in a second individual session after their participation in classes where their knowledge of the SPECSS curriculum was again evaluated by translating native-language items from the curriculum to English. Participants were reimbursed for their individual testing sessions and received classes and SPECSS binders for free.

Participants in the learning group met weekly for 1-h class sessions. During each class, there were two lead teachers, with at least one speaking participants' native language (e.g., an
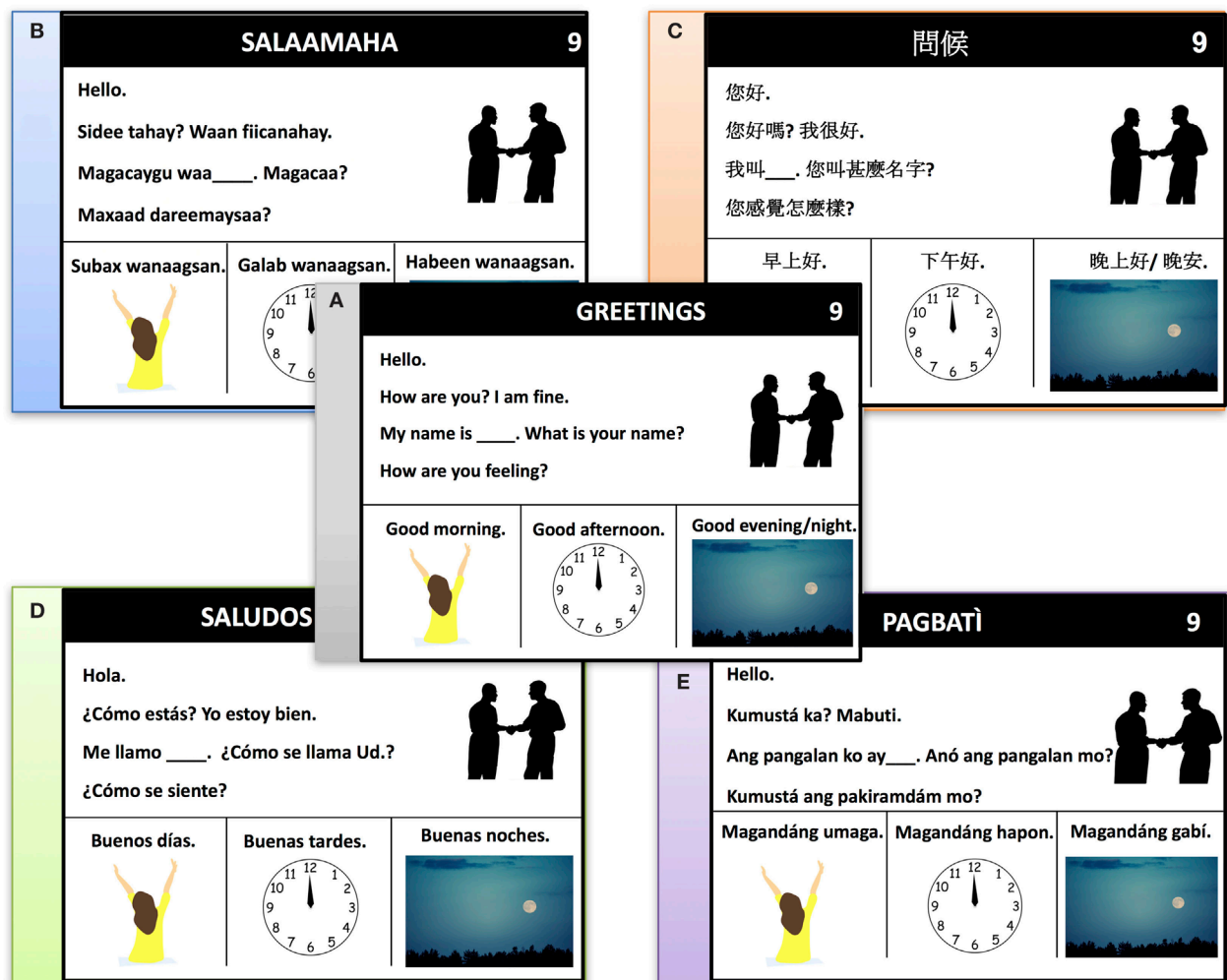
FIGURE 1 | Sample page from the Specific-Purpose English Communication System for Seniors curriculum binder, including the English target page **(A)** together with translation pages in Somali **(B)**, Mandarin **(C)**, Spanish **(D)**, and Tagalog **(E)**. In the curriculum binders, the relevant translation page is printed on the flip side of the English target page, with binders available in English–Somali, English–Mandarin, English–Spanish, and English–Tagalog.

English-speaker and a Somali-English bilingual speaker). In addition, teaching facilitators sat with participants to allow individual practice and provide feedback throughout the session. The teacher-to-student ratio ranged from 1:1 to 1:3. In the first learner cohorts (Mandarin, Spanish, and Tagalog speakers), facilitators sat next to language learners around a square table, with lead teachers at the front. In the Somali cohort, where tables were arranged in a horseshoe shape, facilitators and learners shared a table and faced each other. The lead teacher stood at the front, and presented corresponding content from the SPECSS binder with a screen projector.

Each class session began with lead teachers introducing topics to be covered that day. Next, they presented target words and phrases from the curriculum by saying them in the learners' native language, followed by the English translation. Then, the learners were asked to repeat the English target words and phrases as a group and individually while following along on their binders.

After multiple repetitions of the English words and phrases, participants were given the opportunity to produce the English items after verbal presentation of native-language equivalents to strengthen independent ability to translate targets. Group activities were also employed to practice the novel targets, including dialogs, Bingo, using a map to practice giving directions, etc. Finally, the material was reviewed by asking related conversational questions such as "When is your birthday?" or "How are you feeling today?" The same team of teachers and facilitators taught all classes for each language cohort, allowing for continuity and repeated practice of materials across sessions. Learners participated in an average 11.8 classes based on their availability (SD = 4.5, range: 7–21) and were encouraged to use and practice with their binder outside of classes. For learners in our current study, number of classes attended did not significantly correlate with learning success (newly learned items: $r = 0.14$, $p > 0.5$; items never learned: $r = -0.27$, $p > 0.1$).

Coding and Analyses

Reference Group Data

Montreal Cognitive Assessment

The four subtests from the MoCA (Nasreddine et al., 2005) included in analyses were attention (forward and backward digit span), orientation (time, date, and place), memory (delayed recall), and naming. The total number of points that participants could earn on the *Orientation* subtest was 6. One point was given for each item correctly answered on the orientation subtest: day of the week, month, year, place (name of clinic or office), and city. For the final point, participants who were administered the full MoCA had to name the exact date (e.g., "the 1st" for January 1). On the MoCA-B version, participants had to provide the time. A response that fell within 2 h was accepted.

The *forward and backward digit span* subtest from the MoCA was administered in addition to the MoCA-B. A total of two points could be earned on this subtest. One point each was given for the forward and backward sequence repeated correctly. On the *Memory* (*delayed recall*) subtest, participants were orally instructed to recall five words dictated by research assistants. After dictation, participants were asked to immediately recall the five words. At the end of the test, participants were asked to recall the five words given to them earlier in no specific order required. A total of five points could be earned, with one point given for each target item recalled without any cues. On the MoCA version administered to Mandarin, Spanish, and Tagalog speakers, three animal *naming* cues were provided; on the MoCA-B administered to Somali speakers, four animal cues were provided. Therefore, naming accuracy is reported as a percentage in **Table 1**.

Verbal Fluency

For both animal and grocery tasks, one point was given for each word that was correctly named within its respective category. Repeated words (perseverations) and words that did not match the category cue were not given a point. Synonyms were counted as perseverations (e.g., *papa dulce* and *camote* in Spanish were counted as one item). Further, male and female equivalents of animals were counted as separate items if the phonological form differed by more than one phoneme (e.g., *vaca/toro* counted as two items but *chivo/chiva* counted as one). Participant responses were transcribed on the spot and audio recordings were obtained when permitted. For 42.0% of the data, verbal fluency responses were checked against audio recordings to establish reliability, and reliability was 95.6%.

Similarity to English of Participants' Spoken Languages

To examine the extent to which participants used previous language knowledge to scaffold English acquisition, the similarity to English was coded for languages that participants reported knowledge of on the LEAP-Q, see **Table 2**. Similarity scores were assigned to languages based on their historical similarity to English. A five-point rating system was employed. A score of 1 was assigned to languages that are not Indo-European and have few English borrowings (e.g., Mandarin). A score of 2 was given to languages that are not Indo-European but have some English borrowings. For example, an estimated 10% of the Japanese lexicon consists of words borrowed from English

(McKenzie, 2010). Similarly, Somali has borrowings from English and Italian linked to European colonization (Somali, 2017); see **Table 2** notes for details on other languages. Relative to non-Indo-European languages such as Somali and Japanese, other non-Indo-European languages have an even higher percentage of loan words. A score of 3 was given to languages that are not Indo-European but have substantial English or Spanish influence and borrowings. We categorized most of the reported Philippine languages in this way (e.g., Tagalog) because of evidence of heavy lexical influence of English (Rubino, 2001; Bernardo, 2004) and Spanish (Lipski and Mühlhäusler, 1996; Rubino, 1997; Stolz, 2006; Mattes, 2014). Specifically, English is the default language for many areas of industry (Bernardo, 2004) and has influenced the transformation of formal Tagalog terms into new lexical items (Bautista, 2004). Finally, a score of 4 was assigned to Indo-European languages outside of the Germanic or Romance language families (e.g., Russian) and a score of 5 for languages from within the Germanic or Romance language families (e.g., Spanish). A comparable English-similarity scale was derived by the US State Department Foreign Service Institute based on learning data (language difficulty scale, e.g., Thompson, 1996; Tschirner and Heilenman, 1998). Similarity scores were averaged across participants' languages, yielding one linguistic similarity score indexing the potential for cross-linguistic influence.

SPECSS Learning Data

To focus learning gains on success with functional communication, participants' accuracy in translating native-language words or phrases into English was coded in terms of the semantic content successfully communicated instead of exact words or grammaticality. Pre- and post-learning data were coded on the following scale: 4 = participant did not know the item and gave no response; 3 = participant failed to get their message across but made an attempt (e.g., saying "Saturday" for "Sunday" or "help" for "nurse"); 2 = participant got part of their message across (e.g., saying "money bus" for "bus fare" or "back down" for "lower back"); 1 = participant fully got their message across. Learner responses were divided into four categories based on this scale: (1) *Items that learners knew both pre- and post-curriculum* (*i.e.*, *successfully maintained knowledge across the learning interval*): Items coded as "1" or "2" during both pre- and post-testing; (2) *Items that learners did not know either pre- or post-curriculum*: Items coded as "3" or "4" during both pre- and post-testing; (3) *Items that learners knew pre- but not post-curriculum* (*i.e.*, *forgot*): Items coded as "1" or "2" at pre-test and "3" or "4" at post-test; and (4) *Items that learners knew post- but not pre-curriculum*, (*i.e.*, *items that they newly learned*): Items coded as "3" or "4" at pre-test and "1" or "2" at post-test. For pre- and post-intervention coding of the curriculum, 21% (4 of 19) of participants coded were reviewed by two additional trained researchers, and where disagreement was found, three additional researchers made a final decision by consensus. The average reliability score was 95.6%.

Analyses

To reduce the number of variables included in analyses, principal component, and correlation analyses were first conducted

to identify similar variables. Such variables were combined into cognitive and proficiency indexes by adding their respective z-scores. Specifically, self-reported speaking and comprehension skills correlated both within L1 ($r = 0.74$, $p < 0.001$) and English ($r = 0.89$, $p < 0.001$). Further, animal and grocery verbal fluency scores also correlated within L1 ($r = 0.40$, $p = 0.003$) and English ($r = 0.78$, $p < 0.001$). Therefore, composite self-reported speaking/understanding scores and verbal fluency scores were derived for each language. Only digit span and orientation subtests of the MoCA were found to correlate ($r = 0.29$, $p = 0.017$) and only delayed recall and naming scores were found to correlate ($r = 0.38$, $p = 0.002$), with one component including positive loadings for all four subtests (digit span: 0.49, orientation: 0.55, delayed recall: 0.74, naming: 0.64; eigenvalue = 1.50), and one component including positive loadings for digit span and orientation with negative loadings for delayed recall and naming (digit span: 0.64, orientation: 0.57, delayed recall: −0.38, naming: −0.55; eigenvalue = 1.19). Therefore, composite digit span/orientation and memory/naming scores were derived.

To examine cognitive contributors to language skills and the relation between English and L1 skills, regression analyses were conducted across the full sample of 53 participants. First, to examine effects of cognitive aging across tasks, multivariate regression analyses were conducted with age as a predictor and skills in participants' native language (self-reported proficiency; verbal fluency), English, and cognitive performance (digit span/orientation, memory/naming) as dependent measures. Next, to select the best L1 and cognitive predictors of English skills, regression analyses were conducted. To eliminate variables that were not unique predictors relative to other measures, predictor variables were entered and eliminated from regression models in an iterative backward manner, with the criterion for removal being $p \geq 0.1$. In stepwise regressions, backward entry of variables is preferable to forward entry because the latter is at greater risk for Type II error (Field, 2009, p. 213).

When analyses were conducted on the 19 older adults who participated in our English classes, z-score based scores that were derived in the context of the larger reference group were used to maintain more standardized self-rated proficiency, verbal fluency, and cognitive performance scores. Planned correlation analyses were conducted examining how learners' stable knowledge (items they knew both before and after participating in our English classes) and the newly learned items they had acquired related to linguistic and cognitive predictors that had been identified in the reference group. These predictors included the mean similarity of participants' other languages to English, their exposure to English, as well as composite L1 proficiency and English proficiency scores that were derived by adding participants' z-scores for self-reported speaking and comprehension skills and their z-scores of verbal fluency skills. To alleviate the risk of Type I error due to multiple correlations, confidence intervals for each significant correlation were bootstrapped in SPSS using the bias corrected accelerated 95% confidence interval option, and only significant correlations whose lower bound demonstrated at least a small effect ($r = 0.1$, Cohen, 1988) were interpreted.

RESULTS

Cognitive Predictors of Language Proficiency

Multivariate regression analyses were conducted to examine age-related changes in participants' native language, English, and cognitive skills. Age was found to significantly predict L1 skills, $F(2, 49) = 3.8$, $p = 0.029$, $\eta_p^2 = 0.13$, with increased age significantly predicting L1 speaking/comprehension, beta = −0.1, $t = −2.78$, $p = 0.008$, but not L1 verbal fluency, beta = 0.001, $t = 0.03$, $p > 0.1$. Age did not significantly predict English skills, $F(2, 48) = 1.39$, $p > 0.1$, $\eta_p^2 = 0.06$, or cognitive skills, $F(2, 50) = 1.79$, $p > 0.1$, $\eta_p^2 = 0.07$. Thus, while older adults in our sample were more likely to report lower L1 skills (see **Figure 2A**), no effects of age were observed in L1 verbal fluency, English proficiency, and cognitive skills.
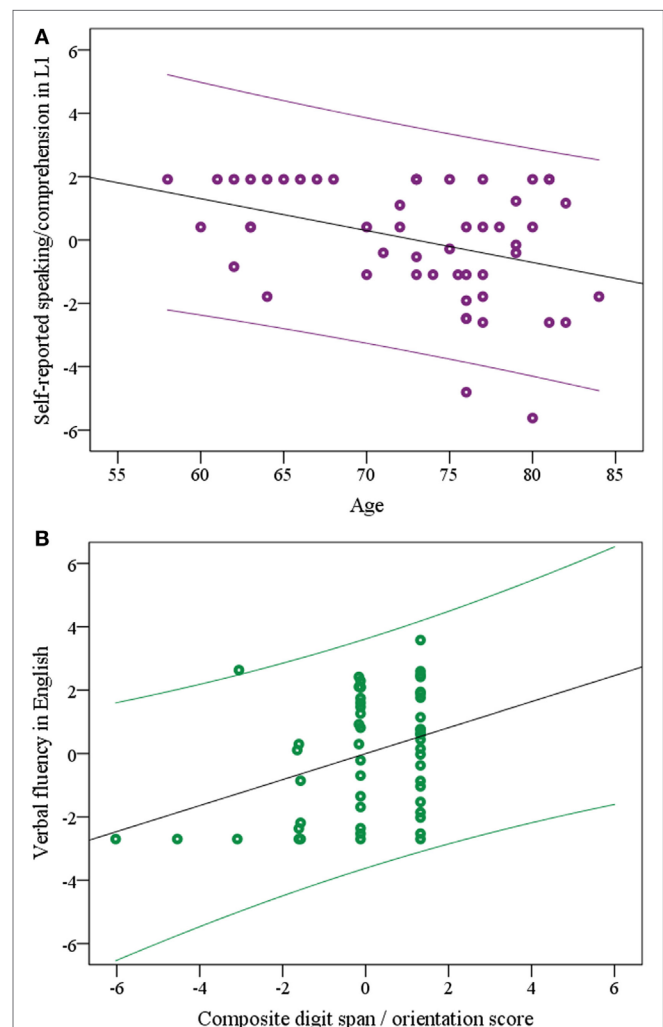

FIGURE 2 | Cognitive predictors of language proficiency: age as a predictor of self-reported speaking/comprehension in L1 **(A)**; composite digit span/orientation as a predictor of verbal fluency in English **(B)**. Pairwise correlations are plotted with error lines representing 95% confidence intervals.

Next, regression analyses were conducted to identify cognitive skills that might support English proficiency across our participants. Composite English speaking/comprehension was entered as a dependent variable with digit span/orientation and memory/naming scores as predictor variables. No significant model emerged, suggesting that none of the cognitive variables predicted self-perceived English-speaking/comprehension proficiency, $F(1, 49) = 0.83$, $p > 0.1$, $R^2 = 0.02$. When English composite verbal fluency was entered as dependent measure with the same predictors, only digit span/orientation emerged as a significant predictor of English verbal fluency (beta = 0.35, $t = 2.66$, $p = 0.01$), $F(1, 51) = 7.08$, $p = 0.01$, $R^2 = 0.12$, see **Figure 2B**. Thus, digit span/orientation, emerged as a cognitive predictor of English proficiency in the current sample of older adult L2 speakers.

Linguistic Experience Predictors of English Proficiency

Regression analyses were conducted to identify the strongest predictors of English proficiency. First, English composite speaking/comprehension was entered into a backward regression analysis as dependent measure, with L1 composite speaking/comprehension, L1 composite verbal fluency, age of first English exposure, current exposure to English, and mean similarity to English of other languages spoken as predictor variables. A significant model emerged, $F(4, 42) = 9.8$, $p < 0.001$, $R^2 = 0.48$, with self-reported English-speaking/comprehension skills predicted by exposure to English (beta = 0.31, $t = 2.41$, $p = 0.021$), by similarity to English of languages known to the participant (beta = 0.42, $t = 3.18$, $p = 0.003$), by verbal fluency in L1 (beta = −0.27, $t = −2.34$, $p = 0.024$), and by age of first exposure to English (beta = −0.24, $t = −2.10$, $p = 0.041$), see **Figure 3**. Similarly, when composite English verbal fluency was entered as a dependent measure with the same predictor variables, a significant model emerged, $F(4, 43) = 16.40$, $p < 0.001$, $R^2 = 0.60$, with English verbal fluency significantly predicted by mean similarity to English of other languages known to the participant (beta = 0.41, $t = 4.09$, $p < 0.001$), by age of first exposure to English (beta = −0.37, $t = −3.80$, $p < 0.001$), by verbal fluency in L1 (beta = 0.32, $t = 3.13$, $p = 0.003$), and by L1 self-reported speaking/comprehension (beta = −0.20, $t = −2.04$, $p = 0.048$), see **Figure 4**. Thus, for both self-reported and verbal fluency measures, better English skills were associated with higher similarity to English of other languages known to the participants and with earlier first exposure to English. Higher self-reported English skills were related to lower L1 verbal fluency, while higher English verbal fluency was related to higher L1 verbal fluency and lower L1 self-reported skills.

Novel Learning of English

The 19 individuals who participated in classes to improve their English showed an average 33.3% increase in their mastery of functional English skills (SE = 8.3, range: −4 to 155%). This increase constituted gain of an average of 78.6 new words or phrases (SE = 9.3, range: 25–151), with an average 18.1 "forgotten" items per participant that were accurately produced before but not after participating in the English course (SE = 2.8,

range: 4–47), see **Table 3**. This gain in English knowledge was found to be statistically significant, with more items translated successfully from L1 to English per participant after the classes ($M = 262.6$, SE = 17.2) than before ($M = 207.2$, SE = 17.5), $t(55) = −5.233$, $p < 0.001$ (items coded as 1 "the meaning was fully communicated"), and with fewer "I don't know" (coded as 4) responses after the classes ($M = 52.1$, SE = 15.1) than before ($M = 100.7$, SE = 21.4), $t(18) = 5.2$, $p < 0.001$. While not statistically significant, a pattern of more items coded as 2 ("message partially communicated") and fewer items coded as 3 ("communication attempted but unsuccessful") after the classes also suggested gradual learning. Critically, participants gained significantly more novel items from pre- to post-testing than they forgot, $t(18) = 5.8$, $p < 0.001$.

Correlation analyses were conducted to examine to what extent the linguistic and cognitive factors that predicted English performance in our reference group were also associated with the stable knowledge learners displayed across their participation span as well as their number of newly mastered items, see **Table 4**. Learners with higher stable knowledge of the curriculum's content across their pre- and post-SPECSS curriculum sessions also showed higher composite English proficiency scores prior to starting the SPECSS classes, $r(18) = 0.79$, $p < 0.001$, higher performance on the digit span and orientation subtests of the MoCA, $r(18) = 0.67$, $p = 0.002$, see **Figures 5A,B**, as well as earlier ages of English acquisition, $r(18) = −0.55$, $p = 0.014$. Instead, learners who acquired the most new items between pre- and post-SPECSS curriculum sessions were found to have the *least* previous knowledge of English, $r(18) = −0.68$, $p = 0.001$, and the *least* similarity between English and their previously known languages, $r(18) = −0.58$, $p = 0.01$, see **Figures 5C,D**.

DISCUSSION

In the current study, we examined how a relatively low-proficient non-native language (English) would be mastered with increased age, including the roles of cognitive skills and previous linguistic experiences. Further, we asked how cognitive and linguistic factors would influence older learners' success in maintaining what they know and acquiring novel functional English through a specific-purpose English curriculum. In a group of older adults with a variety of language backgrounds, we found that age was not a predictor of English verbal fluency performance, short-term language maintenance or learning, but we identified digit span and orientation as potential cognitive predictors. Further, the influence of previous linguistic experiences on English attainment, short-term maintenance, and learning pointed to the roles of both transfer from previously learned languages and continued exposure to English as key variables.

Cognitive Factors in Older Adults' Ability to Maintain and Learn a Foreign Language

The finding across our overall sample of older adults that individuals' age was not related to their English skills is consistent with previous results, suggesting that age-related declines in language are subtle and not pervasive (e.g., Burke and Peters, 1986;
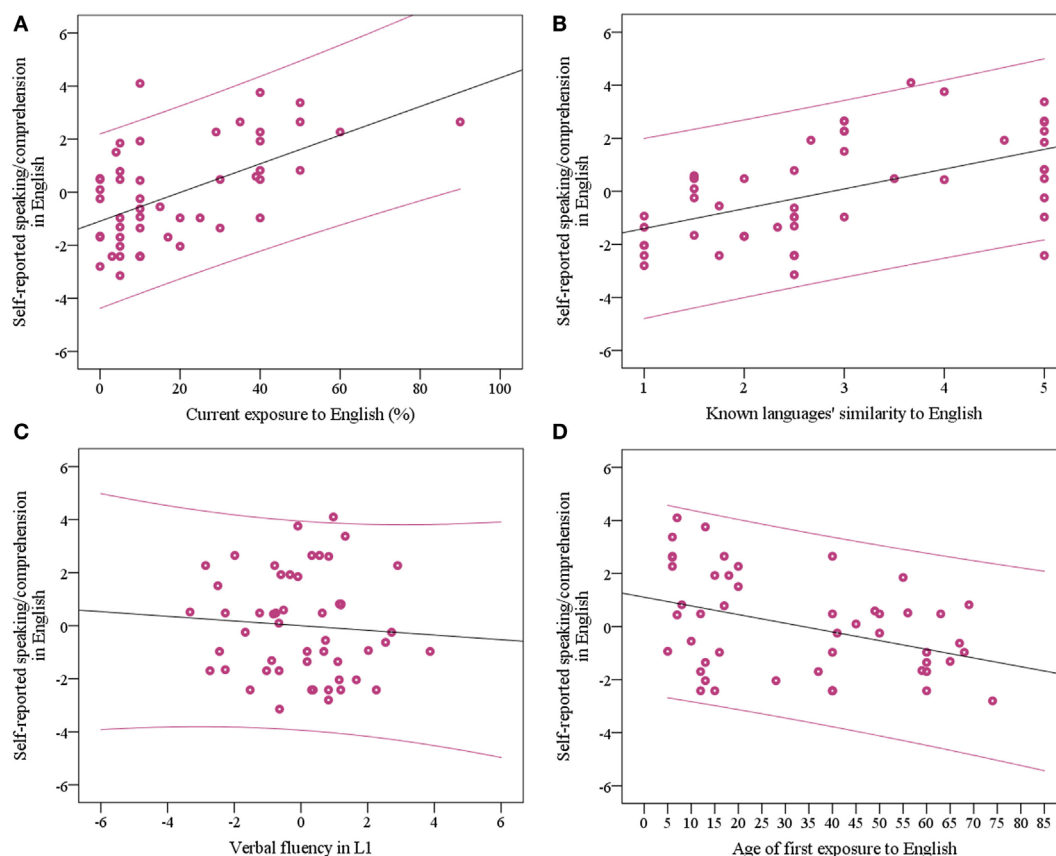
FIGURE 3 | Exposure to English **(A)**, similarity to English of the languages known by participants **(B)**, verbal fluency in L1 **(C)**, and age of first English exposure **(D)** as unique linguistic predictors of the composite self-reported English-speaking and comprehension score. Pairwise correlations are plotted with error lines representing 95% confidence intervals.

Park et al., 2002), with lexical knowledge especially stable (Park et al., 2002), and with age-related decline typically limited to cognitively challenging linguistic contexts (e.g., Kemper, 1986). Interestingly, in the current participant cohort, increased age was associated with lower self-reported speaking and comprehension (but not verbal fluency) in L1. It is possible that this dissociation is tied to L1 attrition, with participants judging their L1 proficiency against a standard of higher skills earlier in life, with lower self-reported ratings further away in time from participants' peak L1 proficiency (i.e., later in life). Instead, age effects on L1 were not captured in *current* verbal fluency performance, suggesting no decline in objective L1 performance. With self-reports found to be reliable but by nature more subjective (e.g., Marian et al., 2007), we believe it best to exercise caution in concluding that marked decline in L1 proficiency is captured in the link between age and self-reported proficiency.

Instead of age, composite scores of digit span and orientation were found to predict English verbal fluency in the reference group as well as short-term maintenance of knowledge in the SPECSS learners. Verbal short-term memory (e.g., Papagno and Vallar, 1995; Kaushanskaya et al., 2011) and attention skills (e.g., Bartolotti et al., 2011) have been linked to the ability to acquire novel vocabulary and process an L2. It is thus consistent

with previous findings that individuals with higher scores on digit span/orientation subtests were more successful at learning English independently prior to our SPECSS curriculum, and that the learners who participated in our SPECSS curriculum better maintained skills across the curriculum. Relatedly, Marcotte and Ansaldo (2014) argued that their monolingual participants' slow initial learning of non-cognates in a novel L2 was linked to age-related declines in the encoding of phonological sequences. Considering that our definition of language maintenance is limited to short-term maintenance across a span of weeks in the current study and given our relatively small sample of SPECSS learners that allowed us to identify this effect in pre- and post-curriculum performance, additional work should be conducted linking attention, and phonological short-term memory to long-term maintenance of low-proficient English in older adults. Indeed, findings linking English performance to cognitive performance in our larger reference group are indicative of the cognitive skills needed to learn and retain a foreign language. If replicated and identified across a longer time window, a link between cognitive performance and continued L2 performance may provide valuable information regarding the continued support and practice resources older adult language learners may need beyond traditional multi-week language courses.
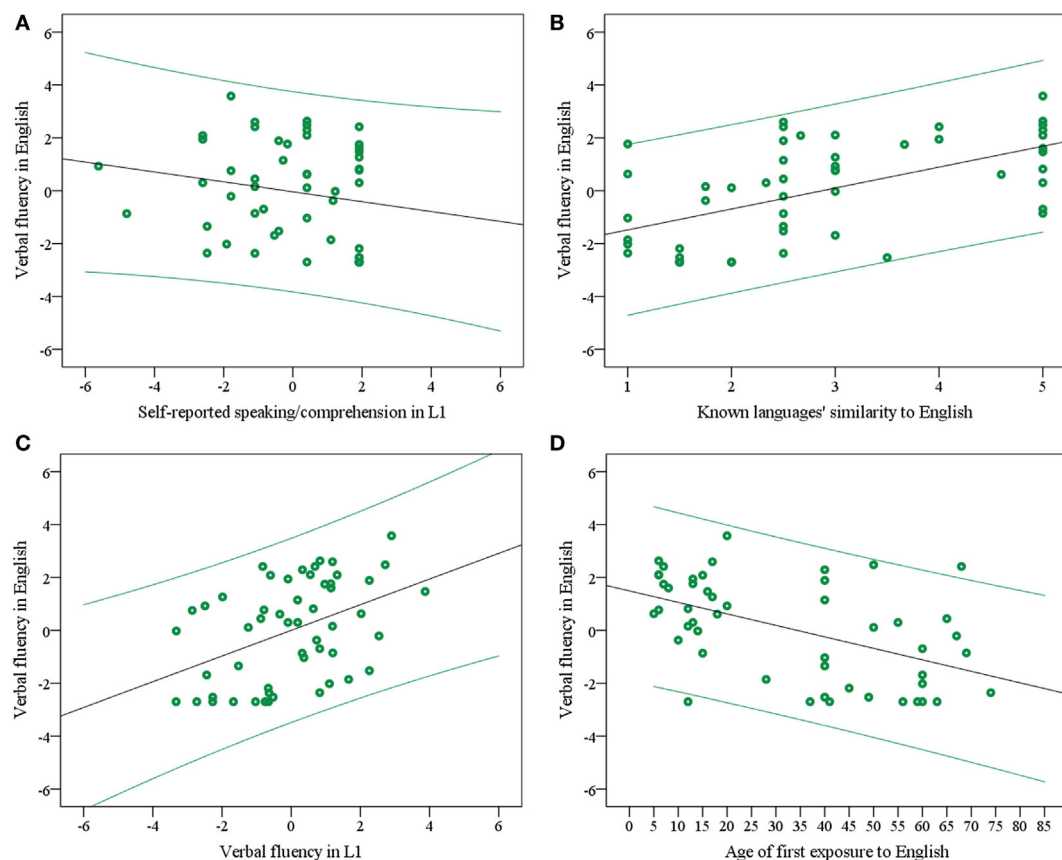
FIGURE 4 | Self-reported speaking/comprehension in L1 **(A)**, similarity to English of the languages known by participants **(B)**, verbal fluency in L1 **(C)**, and age of first exposure to English **(D)** as unique linguistic predictors of English verbal fluency. Pairwise correlations are plotted with error lines representing 95% confidence intervals.

TABLE 3 | Learners' success in acquiring items and phrases through the Specific-Purpose English Communication System for Seniors curriculum.

	Mean	SD	Range
Pre- and post-learning knowledge			
Items communicated successfully (1)			
Pre-learning	207.21	76.47	49–334
Post-learning	262.63	74.99	127–373
Items partially communicated (2)			
Pre-learning	43.11	23.72	11–92
Post-learning	52.63	23.93	21–121
Items not communicated successfully (3)			
Pre-learning	61.47	41.62	5–181
Post-learning	51.21	31.53	10–118
Items participants did not attempt (4)			
Pre-learning	100.68	93.46	7–358
Post-learning	52.05	65.61	1–226
Stable knowledge (known before and after)	230.74	84.89	51–364
Items newly learned (not known before but known after)	78.63	40.73	25–151
Items forgot (known before but not after)	18.05	12.28	4–47
Items never learned (did not know before or after)	80.89	73.68	4–259
Percent increase in knowledge	33.28	36.15	−4.00 to 155

While orientation/digit span composite scores capture cognitive skills that are compelling predictors of language success, other experiential factors have been shown to guide executive function in older adults, most notably educational attainment (e.g., Bosma et al., 2003; Van Hooren et al., 2007). Similarly, aspects of linguistic performance such as verbal fluency have been linked to educational attainment (Van Hooren et al., 2007). Consistently, orientation/digit span performance in our participants correlated with years of formal education, $r(52) = 0.56$, $p < 0.001$, and English verbal fluency also correlated with years of education, $r(51) = 0.36$, $p = 0.01$. In this sense, findings from the current study are also consistent with the premise that older adult language learners with less formal education may be particularly vulnerable in terms of their ability to acquire and maintain novel language knowledge.

Regardless of individual differences in executive function and educational attainment, participants who enrolled in the SPECSS curriculum showed significant *learning effects* at the group level. Learners' ability to acquire *novel* items as part of the SPECSS curriculum was found to be linked to their previous English skills and overall language knowledge rather
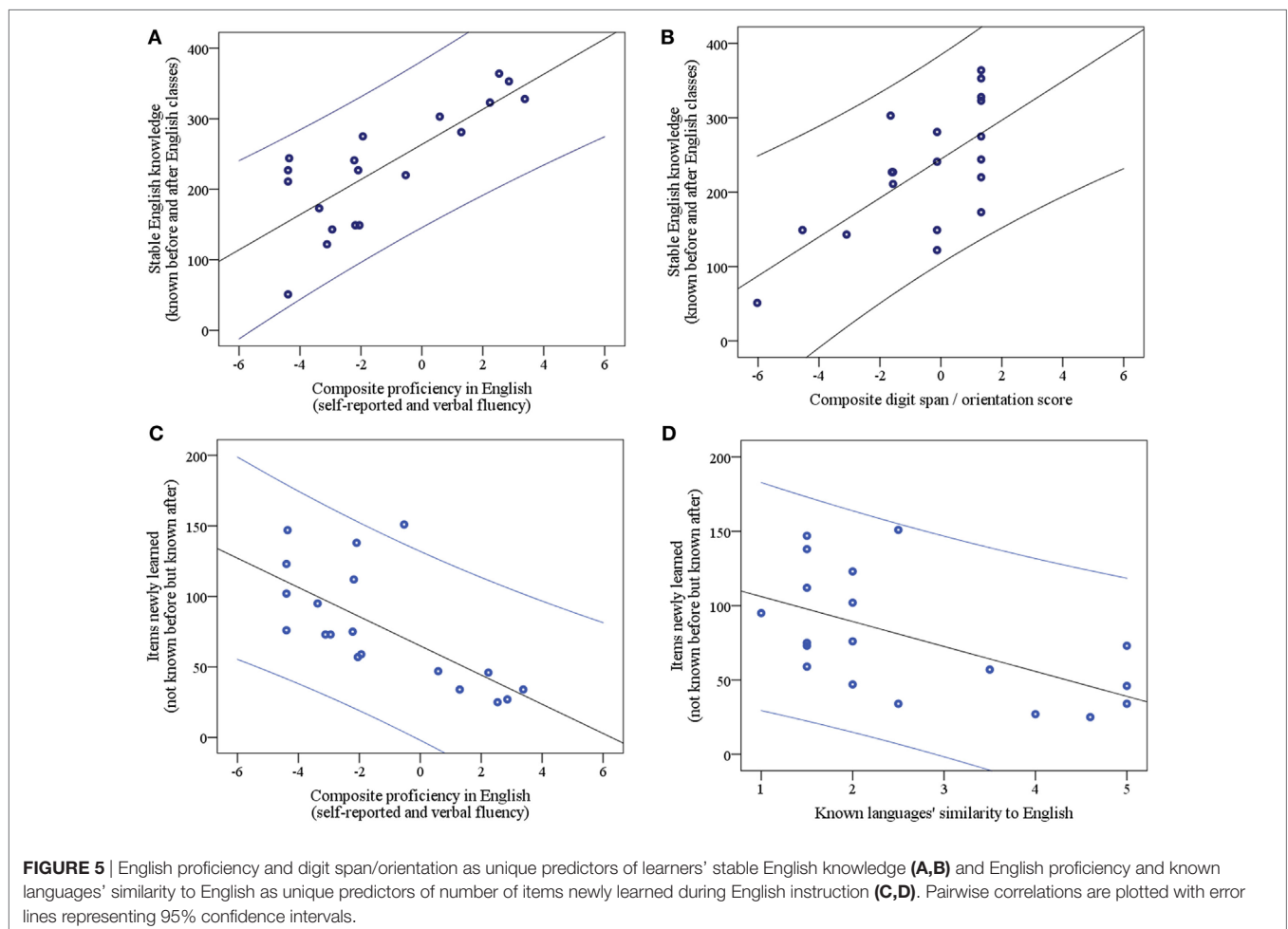
TABLE 4 | Correlations between learners' stable knowledge of the English curriculum, curriculum items newly acquired, as well as key linguistic and cognitive variables.

	Stable English knowledge (known before and after)		Items newly learned (not known before but known after)	
	r (p)	95% CI	*r (p)*	95% CI
Linguistic variables				
Mean similarity to English of known languages	0.34 (ns)		**−0.58 (0.010)**	**−0.79 to −0.29**
Exposure to English	0.37 (ns)		−0.47 (0.042)	−0.73 to −0.06
Age of first English exposure	**−0.55 (0.014)**	**−0.80 to −0.13**	0.42 (ns)	
L1 proficiency[a]	0.29 (ns)		−0.24 (ns)	
English proficiency[a]	**0.79 (<0.001)**	**0.55 to 0.91**	**−0.68 (0.001)**	**−0.89 to −0.35**
Cognitive variables				
Age	−0.14 (ns)		0.23 (ns)	
Years of formal education	0.45 (0.054)	0.08 to 0.72	−0.36 (ns)	
Digit span and orientation	**0.67 (0.002)**	**0.27 to 0.87**	−0.27 (ns)	
Memory and naming	0.07 (ns)		0.02 (ns)	

Confidence intervals are provided to assess reliability of significant correlations (in bold).

CI, confidence interval.

[a]*Combined self-reported and verbal fluency scores. To alleviate risk for Type I error due to multiple comparisons, only significant correlations with a 95% CI lower bound of at least r = 0.1 were interpreted (in bold).*

FIGURE 5 | English proficiency and digit span/orientation as unique predictors of learners' stable English knowledge **(A,B)** and English proficiency and known languages' similarity to English as unique predictors of number of items newly learned during English instruction **(C,D)**. Pairwise correlations are plotted with error lines representing 95% confidence intervals.

than cognitive factors, suggesting that the curriculum was appropriate for older adult learners across a range of cognitive performance levels. This was perhaps the case since many of the cognitive hurdles in adult language learning were addressed as part of the SPECSS curriculum, thus providing scaffolding for learners.

Linguistic Factors in Older Adults' Ability to Maintain and Learn a Foreign Language

In examining linguistic predictors to language attainment, maintenance, and learning in our participants, we examined both acquisition age and exposure to English, and considered participants' overall knowledge of other languages in a combined score indexing similarity to English of other languages spoken. Linguistic predictors of English verbal fluency included stronger L1 verbal fluency, while participants with higher L1 verbal fluency tended to self-report somewhat lower English skills, perhaps because they judged their English against their L1. Together, findings suggest that a combination of linguistic transfer and experience determines older adults' foreign language skills.

Linguistic Transfer

Across our reference group, we found that higher self-reported speaking/comprehension and verbal fluency in English were associated with greater English-similarity of other known languages. These findings are consistent with the prediction that positive transfer from other languages would influence success in acquiring and maintaining English (e.g., MacWhinney, 2012). Findings are also consistent with Marcotte and Ansaldo (2014)'s results that novel words similar to established knowledge (i.e., cognates) are easier to learn than linguistically novel items, particularly in the early stages of learning [also see Bartolotti and Marian (2016) in younger adults]. Since lexico-semantic knowledge has been found to be particularly stable with cognitive aging (Reuter-Lorenz et al., 2000; Park et al., 2002), it is conceivable that older learners are particularly reliant on transfer from previously established lexical knowledge as they acquire a novel language.

Interestingly, data from our subset of English learners who participated in our SPECSS curriculum suggest that those who learned the most novel items during our classes were the learners who had previous languages with the *least* similarity to English. We offer three preliminary explanations for this effect in the spirit of generating hypotheses for future work on mechanisms of language learning in older adults that will hopefully follow these initial findings. First, it is possible that the group who benefited the most from the SPECSS curriculum was comprised of individuals who were the most limited in learning prior to the curriculum. With limited opportunity to transfer knowledge from structurally similar language(s), these individuals may have faced the greatest barriers to independent learning. Such barriers may have been ameliorated by our curriculum by drawing learners' attention to clear functional targets with opportunity for frequent repetition and association with L1 equivalents. As demonstrated by Marcotte and Ansaldo (2014), older adults are capable of acquiring non-cognate knowledge that has low form-relation to a previous language, but it takes considerable effort.

On the flip side, learners who spoke other English-similar languages had previously experienced positive transfer to English during their independent immersion experiences, as evidenced by their higher English skills at the outset of the curriculum. It is possible that, having experienced greater early success in English, these learners were already more functional in their everyday English communications and thus less motivated to acquire novel English knowledge. Alternatively, it is possible that, with more entrenched prior knowledge in English and structurally similar languages, it was in fact more challenging for these learners to acquire additional novel English skills due to competition from stronger languages given cross-linguistic neighbors (e.g., Bartolotti and Marian, 2012) and the expectation of cognate forms (e.g., Siyambalapitiya et al., 2009). Specifically, more entrenched representations become active more rapidly and are more likely to compete with a weaker language, making it potentially more challenging for learners to acquire new items that are similar yet distinct from previous knowledge (e.g., Diependaele, 2012). Research from young adult bilingual vs. monolingual language learners suggests that bilinguals are particularly well-equipped to manage competition from a previous language, a skill that may confer learning advantages relative to monolinguals (Bartolotti and Marian, 2012; Hirosh and Degani, 2017). However, this benefit may be more limited in older multilinguals. It has been suggested that, with cognitive aging, fewer cognitive control resources may be available to resolve cross-linguistic competition of this nature (e.g., Marcotte and Ansaldo, 2014; Blumenfeld et al., 2016b). For example, Marcotte and Ansaldo (2014) found that their younger French-speaking learners recruited cognitive control areas (anterior cingulate cortex and caudate nucleus) while learning novel items in a closely related language, Spanish, and attributed older learners' lack of recruitment of such networks to their slower learning of Spanish targets. Consistent evidence is also available from neuroimaging in older adults that attainment and maintenance of a structurally related second language (Mandarin, with L1 Cantonese) may be cognitively more challenging than attainment and maintenance of a less-related language (English, with L1 Cantonese, Abutalebi et al., 2015).

As an alternative to the above explanations of learning effects, it is possible that, since individuals who learned the most items through the SPECSS curriculum knew the least English at baseline, they were presented with the most learning opportunities through our classes and study of the SPECSS binder. In contrast, new learning opportunities were more limited for individuals who had already established a level of functional English knowledge. If such a possible "ceiling effect" were to underlie the current findings, then it could be predicted that the correlation between number of novel items learned and previous English knowledge would weaken if the number of items participants never learned were accounted for (with fewer items never learned for the most English-proficient individuals, see **Table 3**). Instead, when the number of items never learned was controlled for, the correlation between items newly learned and previous English proficiency became stronger, r (16) $= -0.75$, $p < 0.001$. Given this *post hoc* finding, and given that only 5 of our English learners had less than 30 items that they never learned, we believe that ceiling effects cannot account for the current findings.

Finally, it must be noted that while all participants in our learning group continued to master functional English, they had reported a wide age range of first exposure to English (7–63 years), with best pre-curriculum English attainment outcomes for learners with earlier exposure to English. Therefore, we cannot make conclusions about the age of most efficient language transfer in adult learners. It is conceivable that the most successful

learners had benefited from positive transfer of knowledge at a time in middle adulthood when such transfer was cognitively more efficient, with cognitive control mechanisms more available to mute activation of "false friend" representations that were form-similar yet non-equivalent across languages. Alternatively, in the current age group, it is possible that learners who also knew languages that *differed* from English, were at an advantage in learning novel English because they could globally inhibit these languages (Hirosh and Degani, 2017), a task that may have been cognitively less costly. To our knowledge, no research is currently available examining the success of language transfer across age groups. The possibility of limited benefits in language transfer for older learners warrants additional research.

Continued Exposure to English

In addition to the benefits of language transfer identified in the attainment and maintenance of English, exposure to English emerged as a predictor of English proficiency across our overall sample, consistent with previous findings in younger adults[2] (e.g., Marian et al., 2007; Linck et al., 2009) and older adults (Barresi et al., 1998; Nanchen et al., 2017). It is possible that, in learners who cannot engage efficient cognitive control skills to ameliorate interference from other languages, establishment of language-specific resonance through continued immersion is especially critical in the language acquisition process. Once interference from other languages is reduced, language-specific knowledge can be acquired with a reduced risk of negative transfer and with minimized competition from translation equivalents. It is thus possible that continued exposure to the new language becomes even more critical in older adult learners than it is in younger learners. Findings from Marcotte and Ansaldo (2014) are consistent with this claim, given the slower learning curve of older individuals in their study. Relatedly, it is possible that the need for continued exposure in determining learning success may interact with individual differences in cognitive skills.

Limitations of the Current Study and Future Directions

The current study suggests that older adult learners can make significant functional English gains within a short time in a structured curriculum such as SPECSS, and findings across our overall sample of older adult L2 users point to linguistic and cognitive predictors of L2 proficiency. Given our relatively small subset of SPECSS learners ($n = 19$), additional work is needed to replicate findings of novel learning in older adults, especially given wide confidence intervals observed together with reported correlations; and to identify the specific curriculum components that drive learner success. For example, number of classes attended did not correlate with learning success in our current initial study and we assume that learning may have been based in part on the extent to which participants reviewed their binders outside of class, integrated them into daily interactions, and were

willing to seek out English communication partners outside of the classroom [e.g., see Verga and Kotz (2013) for a call to examine social aspects of L2 acquisition in adults]. We believe that a full understanding of English learning success will rely on further study of these independent learning and social contributors.

In addition, the influence of cross-linguistic similarity on L2 proficiency and novel learning in older adults can be extended to the orthographic level. Given the wide range of reported and observed reading skills in the current participants, their self-reported reading proficiency was not included in analyses because it could not be considered an indicator of their shared core language proficiency. The wide range of reading skills could be tied to years of formal education. For example, within the group of learners, years of education related to both self-reported L1 reading skills, $r(18) = 0.72$, $p = 0.001$, and their English reading skills, $r(18) = 0.46$, $p = 0.05$, with reading skills correlated across the two languages, $r(18) = 0.50$, $p = 0.03$. It has been suggested that orthography can provide significant support in adults' acquisition of foreign languages because the additional modality reinforces new phonological representations, thus creating resonance and overall strengthening of representations (Keshavarz and Astaneh, 2004; MacWhinney, 2012). Indeed, post hoc analyses in the current learners suggested that self-reported English reading skills were tied to greater stable curriculum knowledge, $r(18) = 0.73$, $p < 0.001$. This pattern is consistent with the possibility that written text can further amplify adult learners' ability to specify, consolidate, and maintain novel language representations. It is thus likely that fluent adult readers are provided a critical tool for independent language learning and for continued language maintenance. In particular, it is possible that additional variability exists in our sample based on the nature of speakers' other known orthographies. Specifically, part of the positive transfer from similar languages to English that we observed in the overall group may stem from abilities with an orthography that is similar to English. For example, Koda (1996) suggests that the sound-to-symbol mappings and the nature of orthographic units in L1 may influence L2 reading and Holm and Dodd (1996) found that learners of English were more efficient readers if they had previously learned another alphabetic orthography.

Finally, since learners were part of a classroom setting, with one to three learners per teacher, learners likely did not receive equivalent amounts of attention even though an effort was made to provide one-on-one support. Specifically, learners with the lowest initial language skills may have inadvertently received more attention from instructors and may have learned more for this reason. In addition, instructors observed in retrospect that those with the least English knowledge may have sought out help the most consistently during class sessions. While, in naturalistic teaching settings, such variability in learning support is inevitable, follow-up research in more controlled experimental settings can be conducted to replicate the current findings.

Summary, Future Directions, and Conclusion

In the current study, we identified cognitive and experience-based predictors of English attainment, maintenance, and learning in a multilingual group of older adults with various language backgrounds and with low-English skills. Phonological memory and

---

[2]Robinson Anthony, J. J. D., and Blumenfeld, H.K. Language dominance is predictive of cognate effects and inhibitory control in young adult bilinguals (under review).

orientation, as well as similarity to English of previously known languages and experience with English, emerged as primary predictors of English attainment. Further, preliminary learning data from a service-based intervention suggest that older learners confronted with the most hurdles to independent language learning may benefit the most, and are able to acquire functional novel language, in a highly scaffolded learning context. The current findings can serve in generating hypotheses on determinants of preparedness for language learning in older adults. For example, while transfer from similar languages seemed to be a primary predictor of independent learning success, it is unclear whether success in transferring linguistic knowledge to novel contexts is in itself constrained by cognitive aging. Specifically, when compared with similar learners in middle adulthood, it is possible that older adult learners are less able to identify and minimize negative transfer when learning a language similar to previously known languages. Follow-up research is warranted to directly examine this possibility.

The current findings are also useful in identifying key elements to develop successful language learning curricula for older adults. For example, such elements include an awareness of learners' previous linguistic experiences, including the potential for transfer from other languages and their cognitive skills related to attention, as well as scaffolding through visually based materials that may compensate for cognitive hurdles to learning. In addition, the format of the SPECSS curriculum as highly functional and portable may allow learners to practice and integrate knowledge in the context of daily routines. The benefit to learning may be that material is encoded in contexts similar to where it will be retrieved, thus ensuring learners the support of context-dependent memory (e.g., Marian and Kaushanskaya, 2011). In terms of use, even learners who do not fully master material may carry binders and point to targets in communication settings such as visits with healthcare providers, following the model of alternative augmentative communication devices sometimes used by individuals with verbal communication challenges (e.g., Fried-Oken et al., 2011). Further anecdotal feedback from participants suggests that the bi-directional nature of the curriculum, with English targets and native-language translations present, may facilitate intergenerational communication and learning, giving English-speaking younger family members access to an older family member's L1. Finally, the presence of text with images and auditory repetition during classes may promote English literacy. We believe these functional-social aspects of the SPECSS materials have the potential to provide the scaffolding for language learning and communication needed by many older adults, and future work can examine these aspects of the curriculum.

In the examination of learning mechanisms, additional research is needed to examine how the apparent cumulative benefit from multiple previously known languages in adult learning may relate to bilingual advantages (or lack thereof) identified in other contexts. Previous findings suggest that bilingual learning advantages may be domain-specific and limited to linguistic context that had previously been encountered by the individual (e.g., Kaushanskaya and Rechtzigel, 2012; Antoniou et al., 2014; Blumenfeld and Adams, 2014; Hirosh and Degani, 2017). The current findings are consistent with this literature. However,

interestingly, the cognitive skills that were identified as potential predictors of independent language learning success (digit span and orientation) have also been identified as potential cognitive consequences of long-term bilingualism in older adults (e.g., Kavé et al., 2008). This leaves open the possibility for a somewhat broader maintained ability for language learning in bilingual and multilingual older adults (e.g., Antoniou et al., 2013). It remains an unanswered question whether such bilingual learning advantages extend across the adult lifespan and how they interact with language transfer phenomena.

In general, relatively little work is currently available examining language learning success in older adults, particularly with a view on the previous linguistic and cognitive experiences of such learners (e.g., Antoniou et al., 2013; Marcotte and Ansaldo, 2014). This line of research can provide new insights on the nature and extent of experience-induced plasticity. This knowledge, in turn, is of theoretical value in understanding mechanisms and consequences of learning. It also has tremendous applied potential in a world where many older adults must continue to engage in language learning and where learning success is frequently tied to individuals' ability to navigate their environment. Understanding older language learners' cognitive and experiential strengths and vulnerabilities can lead to the development of learning programs tailored to this population. While we see the current study as valuable in establishing general patterns, generating hypotheses, and validating language learning resources for older adults, we also acknowledge that experimentally more controlled research is needed to confirm and extend findings.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of San Diego State University's Institutional Review Board. The protocol was approved by the San Diego State University Internal Review Board. Written informed consent was obtained from all participants.

AUTHOR CONTRIBUTIONS

HB is responsible for the conception and design of the study, with input from SQ and CA. HB is responsible for data analysis, and HB, SQ, CA, and SR are responsible for data acquisition and interpretation, drafting of the work, revising it critically for intellectual content, and final approval of the version to be published. HB, SQ, CA, and SR are in agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

ACKNOWLEDGMENTS

REFERENCES

Abutalebi, J., Canini, M., Della Rosa, P. A., Green, D. W., and Weekes, B. S. (2015). The neuroprotective effects of bilingualism upon the inferior parietal lobule: a structural neuroimaging study in aging Chinese bilinguals. *J. Neurolinguistics* 33, 3–13. doi:10.1016/j.jneuroling.2014.09.008

Amer, T., Campbell, K. L., and Hasher, L. (2016). Cognitive control as a double-edged sword. *Trends Cogn. Sci.* 20, 905–915. doi:10.1016/j.tics.2016.10.002

Antoniou, M., Gunasekera, G., and Wong, P. M. (2013). Foreign language training as cognitive therapy for age-related cognitive decline: a hypothesis for future research. *Neurosci. Biobehav. Rev.* 37, 2689–2698. doi:10.1016/j.neubiorev.2013.09.004

Antoniou, M., Liang, E., Ettlinger, M., and Wong, P. (2014). The bilingual advantage in phonetic learning. *Biling. Lang. Cogn.* 18, 283–295. doi:10.1017/s1366728914000777

Bak, T. H., Long, M. R., Vega-Mendoza, M., and Sorace, A. (2016). Novelty, challenge, and practice: the impact of intensive language learning on attentional functions. *PLoS ONE* 11:e0153485. doi:10.1371/journal.pone.0153485

Barresi, B. A., Obler, L. K., Au, R., and Albert, M. L. (1998). "Language-related factors influencing naming in adulthood," in *Language and Communication in Old Age: Multidisciplinary Perspectives*, ed. H. E. Hamilton (New York, NY: Garland Publishing), 77–90.

Barreveld, D. J. (2001). *Terrorism in the Philippines: The Bloody Trail of Abu Sayyaf, Bin Laden's East Asian Connection*. Lincoln, NE: Writer's Club Press.

Bartolotti, J., and Marian, V. (2012). Language learning and control in monolinguals and bilinguals. *Cogn. Sci.* 36, 1129–1147. doi:10.1111/j.1551-6709.2012.01243.x

Bartolotti, J., and Marian, V. (2016). Bilinguals' existing languages benefit vocabulary learning in a third language. *Lang. Learn.* 67, 110–140. doi:10.1111/lang.12200

Bartolotti, J., Marian, V., Schroeder, S. R., and Shook, A. (2011). Bilingualism and inhibitory control influence statistical learning of novel word forms. *Front. Psychol.* 2:324. doi:10.3389/fpsyg.2011.00324

Bautista, M. L. S. (2004). Tagalog-English code switching as a mode of discourse. *Asia Pacific Educ. Rev* 5, 226–233. doi:10.1007/BF03024960

Bautista, M. L., and Bolton, K. (eds) (2008). *Philippine English: Linguistic and Literary*. Hong Kong, China: Hong Kong University Press.

Bernardo, A. B. I. (2004). McKinley's questionable bequest: over 100 years of English in Philippine education. *World Englishes* 23, 17–31. doi:10.1111/j.1467-971x.2004.00332.x

Bialystok, E., Craik, F. I., and Luk, G. (2008). Lexical access in bilinguals: effects of vocabulary size and executive control. *J. Neurolinguistics* 21, 522–538. doi:10.1016/j.jneuroling.2007.07.001

Blumenfeld, H. K. (2012). "Cognitive aging," in *Routledge Encyclopedia of Second Language Acquisition*, ed. P. Robinson (Abingdon: Routledge), 84–88.

Blumenfeld, H. K., and Adams, A. M. (2014). Learning and processing of nonverbal symbolic information in bilinguals and monolinguals. *Front. Psychol.* 5:1147. doi:10.3389/fpsyg.2014.01147

Blumenfeld, H. K., Bobb, S. C., and Marian, V. (2016a). The role of language proficiency, cognate status, and word frequency in assessment of verbal fluency in Spanish-English bilinguals. *Int. J. Speech Lang. Pathol.* 18, 190–201. doi:10.3109/17549507.2015.1081288

Blumenfeld, H. K., Schroeder, S. R., Bobb, S. C., Freeman, M. R., and Marian, V. (2016b). Auditory word recognition across the lifespan: links between linguistic and non-linguistic inhibitory control in bilinguals and monolinguals. *Linguist. Approaches Biling.* 6, 119–146. doi:10.1075/lab.14030.blu

Bosma, H., Van Boxtel, M. P., Ponds, R., Houx, P., and Jolles, J. (2003). Education and age-related cognitive decline: the contribution of mental workload. *Educ. Gerontol.* 29, 1–9. doi:10.1080/03601270390157024

Burke, D. M., and Peters, L. (1986). Word associations in old age: evidence for consistency in semantic encoding during adulthood. *Psychol. Aging* 1, 283–292. doi:10.1037/0882-7974.1.4.283

Clark, L. J., Gatz, M., Zheng, L., Chen, Y. L., McCleary, C., and Mack, W. J. (2009). Longitudinal verbal fluency in normal aging, preclinical and prevalent Alzheimer disease. *Am. J. Alzheimers Dis. Other Demen.* 24, 461–468. doi:10.1177/1533317509345154

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd Edn. New York, NY: Lawrence Erlbaum Associates.

Costa, A., Santesteban, M., and Caño, A. (2005). On the facilitatory effects of cognate words in bilingual speech production. *Brain Lang.* 94, 94–103. doi:10.1016/j.bandl.2004.12.002

Craik, F. I. M. (2002). Levels of processing: past, present and future? *Memory* 10, 305–318. doi:10.1080/09658210244000135

De Groot, A. M. B. (1992). "Bilingual lexical representation: a closer look at conceptual representations," in *Orthography, Phonology, Morphology, and Meaning*, eds R. Frost and L. Katz (Amsterdam: Elsevier Science Publishers), 389–412.

Diependaele, K. (2012). The word frequency effect in first- and second-language word recognition: a lexical entrenchment account. *Q. J. Exp. Psychol.* 66, 843–863. doi:10.1080/17470218.2012.720994

Ding, H., and Hargraves, L. (2009). Stress-associated poor health among adult immigrants with a language barrier in the United States. *J. Immigr. Minor. Health* 11, 446–452. doi:10.1007/s10903-008-9200-0

Field, A. (2009). *Discovering Statistics Using SPSS*, 3rd Edn. London: SAGE Publications.

Fried-Oken, M., Beukelman, D. R., and Hux, K. (2011). Current and future AAC research considerations for adults with acquired cognitive and communication impairments. *Assist. Technol.* 24, 56–66. doi:10.1080/10400435.2011.648713

Gollan, T. H., Montoya, R. I., and Werner, G. A. (2002). Semantic and letter fluency in Spanish-English bilinguals. *Neuropsychology* 16, 562–576. doi:10.1037/0894-4105.16.4.562

Gower, R. H. (1952). Swahili borrowings from English. *Africa* 22, 154–157. doi:10.2307/1156242

Gregoire, J., and Van der Linden, M. (1997). Effect of age on forward and backward digit spans. *Aging Neuropsychol. Cogn.* 4, 140–149. doi:10.1080/13825589708256642

Hirosh, Z., and Degani, T. (2017). Direct and indirect effects of multilingualism on novel language learning: an integrative review. *Psychon. Bull. Rev.* 1–25. doi:10.3758/s13423-017-1315-7

Holm, A., and Dodd, B. (1996). The effect of first written language on the acquisition of English literacy. *Cognition* 59, 119–147. doi:10.1016/0010-0277(95)00691-5

Hughes, E., and Tainturier, M. (2015). "The cognate advantage in bilingual aphasia: now you see it, now you don't," in *Frontiers in Psychology Conference Abstract: Academy of Aphasia 53rd Annual Meeting*, Tucson, United States.

Janacsek, K., Fiser, J., and Nemeth, D. (2012). The best time to acquire new skills: age-related differences in implicit sequence learning across the human lifespan. *Dev. Sci.* 15, 496–505. doi:10.1111/j.1467-7687.2012.01150.x

Julayanont, P., Phillips, N., Chertkow, H., and Nasreddine, Z. S. (2012). *The Montreal Cognitive Assessment (MOCA): Concept and Clinical Review*. Available at: http://www.pphp.concordia.ca/fac/phillips/files/Julayanontetal2013MoCAreview.pdf

Julayanont, P., Tangwongchai, S., Hemrungrojn, S., Tunvirachaisakul, C., Phanthumchinda, K., Hongsawat, J., et al. (2015). The Montreal cognitive assessment—basic: a screening tool for mild cognitive impairment in illiterate and low-educated elderly adults. *J. Am. Geriatr. Soc.* 63, 2550–2554. doi:10.1111/jgs.13820

Kaushanskaya, M. (2012). Cognitive mechanisms of word learning in bilingual and monolingual adults: the role of phonological memory. *Biling. Lang. Cogn.* 15, 470–489. doi:10.1017/s1366728911000472

Kaushanskaya, M., Blumenfeld, H. K., and Marian, V. (2011). The relationship between vocabulary and short-term memory measures in monolingual and bilingual speakers. *Int. J. Biling.* 15, 408–425. doi:10.1177/1367006911403201

Kaushanskaya, M., and Rechtzigel, K. (2012). Concreteness effects in bilingual and monolingual word learning. *Psychon. Bull. Rev.* 19, 935–941. doi:10.3758/s13423-012-0271-5

Kavé, G., Eyal, N., Shorek, A., and Cohen-Mansfield, J. (2008). Multilingualism and cognitive state in the oldest old. *Psychol. Aging* 23, 70–78. doi:10.1037/0882-7974.23.1.70

Kemper, S. (1986). Imitation of complex syntactic constructions by elderly adults. *Appl. Psycholinguist.* 7, 277–287. doi:10.1017/s0142716400007578

Keshavarz, M. H., and Astaneh, H. (2004). The impact of bilinguality on the learning of English vocabulary as a foreign language (L3). *Int. J. Biling. Educ. Biling.* 7, 295–302. doi:10.1080/13670050408667814

Koda, K. (1996). L2 word recognition research: a critical review. *Mod. Lang. J.* 80, 450–460. doi:10.1111/j.1540-4781.1996.tb05465.x

Kroll, J. F., and Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *J. Mem. Lang.* 33, 149–174. doi:10.1006/jmla.1994.1008

Lee, C., and Federmeier, K. D. (2011). Differential age effects on lexical ambiguity resolution mechanisms. *Psychophysiology* 48, 960–972. doi:10.1111/j.1469-8986.2010.01158.x

Lemhöfer, K., Dijkstra, T., and Michel, M. (2004). Three languages, one ECHO: cognate effects in trilingual word recognition. *Lang. Cogn. Processes* 19, 585–611. doi:10.1080/01690960444000007

Lillard, A. S., and Erisir, A. (2011). Old dogs learning new tricks: neuroplasticity beyond the juvenile period. *Dev. Rev.* 31, 207–239. doi:10.1016/j.dr.2011.07.008

Linck, J. A., Kroll, J. F., and Sunderman, G. (2009). Losing access to the native language while immersed in a second language: evidence for the role of inhibition in second-language learning. *Psychol. Sci.* 20, 1507–1515. doi:10.1111/j.1467-9280.2009.02480.x

Lipski, J. M. (2013). Remixing a mixed language: the emergence of a new pronominal system in Chabacano (Philippine Creole Spanish). *Int. J. Biling.* 17, 448–478. doi:10.1177/1367006912438302

Lipski, J. M., Mühlhäusler, P., and Duthin, F. (1996). "Metropolitan language (Spanish, including Creoles): The Spanish language in the Pacific area," in *Atlas of Languages of Intercultural Communication in the Pacific, Asia, and the Americas*, eds S. A. Wurm, P. Mühlhäusler, and D. T. Tryon (Berlin and New York, NY: Mouton de Gruyter), 271–298.

Lotto, L., and De Groot, A. M. B. (1998). Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Lang. Learn.* 48, 31–69. doi:10.1111/1467-9922.00032

Lustig, C., Hasher, L., and Zacks, R. T. (2007). "Inhibitory deficit theory: recent developments in a "new view"," in *Inhibition in Cognition*, eds D. S. Gorfein and C. M. MacLeod (Washington, DC: American Psychological Association), 145–162.

McKenzie, R. M. (2010). *The Social Psychology of English As a Global Language: Attitudes, Awareness and Identity in the Japanese Context*. London: Springer.

MacWhinney, B. (2005). "A unified model of language acquisition," in *Handbook of Bilingualism: Psycholinguistic Approaches*, eds J. F. Kroll and A. M. B. De Groot (New York, NY: Oxford University Press), 49–67.

MacWhinney, B. (2012). "The logic of the unified model," in *The Routledge Handbook of Second Language Acquisition*, eds S. Gass and A. Mackey (Abingdon: Routledge), 211–227.

Marcotte, K., and Ansaldo, A. (2014). Age-related behavioural and neuro-functional patterns of second language word learning: different ways of being successful. *Brain Lang.* 135, 9. doi:10.1016/j.bandl.2014.04.004

Marian, V., Blumenfeld, H., and Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *J. Speech Lang. Hear. Res.* 50, 940–967. doi:10.1044/1092-4388(2007/067)

Marian, V., and Kaushanskaya, M. (2011). "Language-dependent memory: insights from bilingualism," in *Relations between Language and Memory. Sabest. Saarbrucker*, ed. C. Zelinsky-Wibbelt (Frankfurt: Peter Lang), 95–120.

Marinova-Todd, S. H., Marshall, D. B., and Snow, C. E. (2000). Three misconceptions about age and L2 learning. *TESOL Q.* 34, 9. doi:10.2307/3588095

Mattes, V. (2014). *Types of Reduplication: A Case Study of Bikol*. Boston: De Gruyter Mouton.

Meyer, A. M., and Federmeier, K. D. (2010). Event-related potentials reveal the effects of aging on meaning selection and revision. *Psychophysiology* 47, 673–686. doi:10.1111/j.1469-8986.2010.00983.x

Miyake, A., and Friedman, N. P. (1998). "Individual differences in second language proficiency: working memory as language aptitude," in *Foreign Language Learning: Psycholinguistic Studies on Training and Retention*, eds A. F. Healy and L. E. Bourne (Mahwah, NJ: Lawrence Erlbaum Associates), 339–364.

Mohr, B., Difrancesco, S., Harrington, K., Evans, S., and Pulvermüller, F. (2014). Changes of right-hemispheric activation after constraint-induced, intensive language action therapy in chronic aphasia: FMRI evidence from auditory semantic processing. *Front. Hum. Neurosci.* 8:919. doi:10.3389/fnhum.2014.00919

Morett, L. M., and MacWhinney, B. (2013). Syntactic transfer in English-speaking Spanish learners. *Biling. Lang. Cogn.* 16, 132–151. doi:10.1017/S1366728912000107

Mui, A., Kang, S., Kang, D., and Domanski, M. (2007). English language proficiency and health-related quality of life among Chinese and Korean immigrant elders. *Health Soc. Work* 32, 119–127. doi:10.1093/hsw/32.2.119

Nanchen, G., Abutalebi, J., Assal, F., Mancheon, M., Demonet, J. F., and Annoni, J. M. (2017). Second language performances in elderly bilinguals and individuals with dementia: the role of L2 immersion. *J. Neurolinguistics* 43, 49–58. doi:10.1016/j.jneuroling.2016.09.004

Nasreddine, Z. S., Phillips, N. A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., et al. (2005). The Montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. *J. Am. Geriatr. Soc.* 53, 695–699. doi:10.1111/j.1532-5415.2005.53221.x

Papagno, C., Valentine, T., and Baddeley, A. (1991). Phonological short-term memory and foreign-langauge vocabulary learning. *J. Mem. Lang.* 30, 331–347. doi:10.1016/0749-596x(91)90040-q

Papagno, C., and Vallar, G. (1995). Verbal short-term memory and vocabulary learning in polyglots. *Q. J. Exp. Psychol. A* 48, 98–107. doi:10.1080/14640749508401378

Park, D., Lautenschlager, G., Hedden, T., Davidson, N., Smith, A., and Smith, P. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychol. Aging* 17, 299–320. doi:10.1037/0882-7974.17.2.299

Ponce, N., Hays, R., and Cunningham, W. (2006). Linguistic disparities in health care access and health status among older adults. *J. Gen. Intern. Med.* 21, 786–791. doi:10.1111/j.1525-1497.2006.00491.x

Portocarrero, J. S., Burright, R. G., and Donovick, P. J. (2007). Vocabulary and verbal fluency of bilingual and monolingual college students. *Arch. Clin. Neuropsychol.* 22, 415–422. doi:10.1016/j.acn.2007.01.015

Potapova, I., Blumenfeld, H. K., and Pruitt-Lord, S. (2016). Cognate identification methods: impacts on the cognate advantage in adult and child Spanish-English bilinguals. *Int. J. Bilingual.* 20, 714–731. doi:10.1177/1367006915586586

Raboyeau, G., Marcotte, K., Adrover-Roig, D., and Ansaldo, A. I. (2010). Brain activation and lexical learning: the impact of learning phase and word type. *Neuroimage* 49, 2850–2861. doi:10.1016/j.neuroimage.2009.10.007

Reuter-Lorenz, P. A., Jonides, J., Smith, E. E., Hartley, A., Miller, A., Marshuetz, C., et al. (2000). Age differences in the frontal lateralization of verbal and spatial working memory revealed by PET. *J. Cogn. Neurosci.* 12, 174–187. doi:10.1162/089892900561814

Roberts, P. M., and Deslauriers, L. (1999). Picture naming of cognate and non-cognate nouns in bilingual aphasia. *J. Commun. Disord.* 32, 1–22. doi:10.1016/s0021-9924(98)00026-4

Rosselli, M., Ardila, A., Araujo, K., Weekes, V. A., Caracciolo, V., Padilla, M., et al. (2000). Verbal fluency and repetition skills in healthy older Spanish-English bilinguals. *Appl. Neuropsychol.* 7, 17–24. doi:10.1207/s15324826an0701_3

Rubino, C. R. G. (1997). *A Reference Grammar of Ilocano (Doctoral Dissertation)*. ProQuest Dissertations Publishing, University of California, Santa Barbara.

Retrieved from ProQuest Dissertations and Theses Database. (UMI No. 9823418).

Rubino, C. R. G. (2001). "A reference grammar of Ilocano," in *Ilocano Dictionary and Grammar: Ilocano-English, English-Ilocano*, ed. C. R. Galvez Rubino (Hawaii: University of Hawaii Press), 28.

Ryan, C. (2013). *Language use in the United States: 2011, American Community Survey Reports*. Available at: http://census.gov/prod/2013pubs/acs-22.pdf

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

Siyambalapitiya, S., Chenery, H. J., and Copland, D. A. (2009). Bilingualism and aging: reversal of the cognate advantage in older bilingual adults. *Appl. Psycholinguist.* 30, 531–554. doi:10.1017/s0142716409090262

Somali. (2017). *Accredited Language Services*. Available at: https://www.alsintl.com/resources/languages/Somali/

Sparks, R., Patton, J., Ganschow, L., and Humbach, N. (2009). Long-term crosslinguistic transfer of skills from L1 to L2. *Lang. Learn.* 59, 203–243. doi:10.1111/j.1467-9922.2009.00504.x

Stolz, T. (2006). "Contact-induced typological change," in *Marginal Linguistic Identities: Studies in Slavic Contact and Borderland Varieties*, eds D. Stern and C. Voss (Wiesbaden: Harrassowitz), 13–30.

Stuart, G., Patel, J., and Bhagrath, N. (2006). Ageing affects conceptual but not perceptual memory processes. *Memory* 14, 345. doi:10.1080/09658210500357059

Thompson, I. (1996). Assessing foreign language skills: data from Russian. *Mod. Lang. J.* 80, 47–65. doi:10.1111/j.1540-4781.1996.tb01137.x

Tschirner, E., and Heilenman, L. K. (1998). Reasonable expectations: oral proficiency goals for intermediate-level students of German. *Mod. Lang. J.* 82, 147–158. doi:10.1111/j.1540-4781.1998.tb01187.x

Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: the declarative/procedural model. *Biling. Lang. Cogn.* 4, 105–122. doi:10.1017/s1366728901000220

Van Hooren, S. A. H., Valentijn, A. M., Bosma, H., Ponds, R. W. H. M., Van Boxtel, M. P. J., and Jolles, J. (2007). Cognitive functioning in healthy older adults aged 64–81: a cohort study into the effects of age, sex, and education. *Aging Neuropsychol. Cogn.* 14, 40–54. doi:10.1080/138255890969483

Verga, L., and Kotz, S. A. (2013). How relevant is social interaction in second language learning? *Front. Hum. Neurosci.* 7:550. doi:10.3389/fnhum.2013.00550

Whitfield, J. A., and Goberman, A. M. (2017). Speech motor sequence learning: acquisition and retention in Parkinson disease and normal aging. *J. Speech Lang. Hear. Res.* 60, 1477–1492. doi:10.1044/2016_jslhr-s-16-0104

Bilingual Contexts Modulate the Inhibitory Control Network

Jing Yang[1], Jianqiao Ye[1], Ruiming Wang[2], Ke Zhou[3,4] and Yan Jing Wu[5]*

[1] Bilingual Cognition and Development Lab, Center for Linguistics and Applied Linguistics, Guangdong University of Foreign Studies, Guangzhou, China, [2] Guangdong Provincial Key Laboratory of Mental Health and Cognitive Science, Center for Studies of Psychological Application, School of Psychology, South China Normal University, Guangzhou, China, [3] College of Psychology and Sociology, Shenzhen University, Shenzhen, China, [4] Shenzhen Key Laboratory of Affective and Social Cognitive Science, Shenzhen University, Shenzhen, China, [5] Faculty of Foreign Languages, Ningbo University, Ningbo, China

The present functional magnetic resonance imaging (fMRI) study investigated influences of language contexts on inhibitory control and the underlying neural processes. Thirty Cantonese–Mandarin–English trilingual speakers, who were highly proficient in Cantonese (L1) and Mandarin (L2), and moderately proficient in English (L3), performed a picture-naming task in three dual-language contexts (L1-L2, L2-L3, and L1-L3). After each of the three naming tasks, participants performed a flanker task, measuring contextual effects on the inhibitory control system. Behavioral results showed a typical flanker effect in the L2-L3 and L1-L3 condition, but not in the L1-L2 condition, which indicates contextual facilitation on inhibitory control performance by the L1-L2 context. Whole brain analysis of the fMRI data acquired during the flanker tasks showed more neural activations in the right prefrontal cortex and subcortical areas in the L2-L3 and L1-L3 condition on one hand as compared to the L1-L2 condition on the other hand, suggesting greater involvement of the cognitive control areas when participants were performing the flanker task in L2-L3 and L1-L3 contexts. Effective connectivity analyses displayed a cortical-subcortical-cerebellar circuitry for inhibitory control in the trilinguals. However, contrary to the right-lateralized network in the L1-L2 condition, functional networks for inhibitory control in the L2-L3 and L1-L3 condition are less integrated and more left-lateralized. These findings provide a novel perspective for investigating the interaction between bilingualism (multilingualism) and inhibitory control by demonstrating instant behavioral effects and neural plasticity as a function of changes in global language contexts.

Keywords: bilingualism, inhibitory control, dual-language contexts, fMRI, effective connectivity

INTRODUCTION

Bilingualism is a form of "mental juggler" (Kroll and Bialystok, 2013), as speaking one language often involves simultaneous access to the non-target language in the brain (Dijkstra and Van Heuven, 1998; Green, 1998; Bialystok, 2007; Thierry and Wu, 2007; Wu and Thierry, 2010, 2017; De Groot, 2012). Therefore, for bilingual speakers, managing two languages requires constantly selecting words in the intended language and suppressing activations of the non-target language, a routine that necessitates the engagement of inhibitory control. As a result, the experience of using multiple languages might enhance bilinguals' performance in non-linguistic domains.

Previous studies have shown that bilinguals are less distracted, as compared to monolinguals, when performing inhibitory control tasks, including the Simon task (Bialystok et al., 2004; Martin-Rhee and Bialystok, 2008), the Stroop task (Bialystok et al., 2008), and the flanker task (Costa et al., 2008). Moreover, interpreting training has been shown to improve inhibitory control processes (Dong and Zhong, 2017).

The past decade has witnessed a dramatic increase in the use of neuroimaging techniques such as functional magnetic resonance imaging (fMRI) to study the neural system of bilingual language control and the effects of second language acquisition on inhibitory control (e.g., Bialystok et al., 2005; Luk et al., 2010; Weissberger et al., 2015). Abutalebi and Green (2007) proposed a brain network for language control during bilingual speech production. The network includes the left prefrontal cortex, anterior cingulate cortex (ACC), basal ganglia, and inferior parietal lobule. Abutalebi and Green's (2007) hypothesis is that this neural network is dedicated to the selection and temporal sequencing of language representations during bilingual word production, and the pipeline works in the following order: The left basal-ganglia and ACC modulate the neural activity levels in the left prefrontal cortex, which influences neural activity in the inferior parietal cortex. Each of these areas has implications in distinct cognitive processes: The prefrontal cortex inhibits the non-target language and corrects errors; the ACC monitors conflicts and detects errors; the basal ganglia, especially the caudate nuclei, supervises the language selection and lexical access; the inferior parietal lobule, as a key region for working memory, serves a goal maintenance function. Abutalebi and Green (2008) further clarified distinct contributions of the left and right supramarginal gyri (SMG) in the inferior parietal lobules: The Left SMG is responsible for bias selection away from the language not in use; on the contrary, the right SMG is responsible for bias selection toward the language in use. This neural network for bilingual language control has been testified in a number of studies. In an fMRI study on German–Italian–English trilinguals, Abutalebi et al. (2013a,b) showed that language-switching directions influenced brain activation levels in the caudate nuclei, while activation levels of the supplementary motor area (SMA)/ACC did not vary as the function of language proficiency, which suggests a domain-general role for SMA/ACC in control tasks. Using meta-analysis approaches, Luk et al. (2012) showed that 10 neuroimaging studies involving language switching reported significant and reliable neural responses in the following brain regions: The left inferior frontal gyrus (IFG), left middle temporal gyrus (MTG), left middle frontal gyrus (MFG), right precentral gyrus, right superior temporal gyrus (STG), midline pre-supplementary motor area, and bilateral caudate nuclei. Taken together, it is clear that the prefrontal cortex and caudate nuclei are highly involved in the regulation of bilingual speeches. These brain areas have also been reported in studies of non-linguistic cognitive tasks (e.g., Luk et al., 2010). However, how these areas are connected with one another as part of the control network remains unclear.

Taking the advantage of the functional connectivity approach, researchers have attempted to identify the interactions between the language control network and the cognitive control network

during L2 acquisition. For example, Ghazi Saidi et al. (2013) in an L2 vocabulary training study showed that the language processing network and cognitive control network were highly integrated at the initial stage of vocabulary learning, but as the learning proceeded and the vocabularies are being consolidated, this integration decreased. Grant et al. (2015) expanded this line of research. Instead of lab-based short-term vocabulary training, they studied neural adaptations in the development of L2 processing by examining a group of classroom Spanish L2 learners who were native English speakers over the course of one academic year. Their results show that with increased L2 experience, the overall activations in the control areas such as the ACC decreases while its connectivities with semantically related regions such as the MTG increases. The authors claim that the ability to utilize cognitive control mechanisms to regulate access to the L2 representations is a more critical issue in the beginning, relative to the latter stage, of L2 acquisition. Taken together, these studies suggest an important role of the cognitive control network in early L2 acquisition.

One possibility is that the high demand and long-term practice of language control, which involves inhibitory control, during in L2 acquisition, that allows bilinguals outperforming monolinguals in several cognitive control tasks. However, participants' background variables, such as socioeconomic status and ethnic origins (Sabbagh et al., 2006; Morton and Harper, 2007; Li et al., 2013), also seem to matter for the cognitive advantage in bilinguals. It is also possible that language processing contexts account for some of the variances (e.g., Wu and Thierry, 2013). Studies of neural plasticity on high temporal scales (Fields, 2005; Delekate et al., 2011; Bercury and Macklin, 2015) support the notion that different global language contexts (single or dual-language contexts) may lead to distinguished neural activation patterns during target word selection (Green, 2011). Green and Abutalebi (2013) proposed an adaptive control hypothesis: Language control processes adapt to the recurrent demands of the interactional context. For example, in a dual-language context, in which both languages are used (but to different speakers), language processing engages the control network comprising bilaterally inferior frontal and parietal cortices, the ACC/pre-SMA, basal ganglia, and thalamus (Abutalebi and Green, 2016). In a dense code-switching context, however, speakers routinely interleave their languages in the course of a single utterance and adapt words from one of language in the context of the other language. The neural network of language control would rely more on a cerebellar-prefrontal connection as compared to the dual-language context because, in a dense code-switching context, language control involves higher demands for opportunistic planning (Abutalebi and Green, 2016).

Although the adaptive control hypothesis is a recent theory on the neural mechanisms of bilingual control, there has been increasing interest in the influence of language contexts on non-linguistic executive functions, such as inhibitory control. In a study using event-related potentials (ERPs), Wu and Thierry (2013) examined effects of immediate changes in language processing contexts on executive function in a group of early Welsh–English bilinguals. The cognitive control performance of

these participants was measured using a modified version of the classic flanker task, in which participants were instructed to press a button to indicate the direction of an arrow presented within an array of flankers (arrows pointing to the same or the opposite direction). Critically, a word is presented before the flanker trial to implicitly prime a language context. The contextual words were either in Welsh (L1), English (L2), or both languages in separated blocks. The results showed higher accuracy rates when bilingual participants performed incongruent trials of the flanker task in the dual-language context as compared to single-language contexts. The P300 amplitude was also reduced in the dual-language, as compared to the single-language context, indicating less flanker interference effect in the Welsh–English context. Therefore, the authors claimed that changes in language processing contexts could modulate non-linguistic cognitive control in bilinguals.

In a further exploration, Liu et al. (2016) examined the effect of language contexts on cognitive control in a group of Chinese–English bilinguals. Unlike the highly proficient Welsh–English bilinguals in Wu and Thierry (2013), participants in Liu et al. (2016) were native speakers of Chinese who have a moderate level of proficiency in English. All participants performed an antisaccade task, which measures response inhibition (or response suppression), interference inhibition (or inhibitory control), and task switching, three key subcomponents of executive functions (Bialystok et al., 2006). Response suppression refers to the ability to withhold an inappropriate response (e.g., triggered by a habitual cue), as is most classically established in the go/no-go paradigm. Inhibitory control refers to the process when multiple sources of information (e.g., the printing color and the word meaning in the classic Stroop paradigm) are competing for attention which needs to be drawn to the target attribute of the stimulus. Task switching refers to the ability to alter between two tasks that require different cognitive processes and responses. The critical difference between response suppression and inhibitory control is that the former taps onto the process of response execution, whereas the latter mainly measures the control of selective attention. In Liu et al. (2016), Chinese–English participants performed an antisaccade task in the pre-test and then complete a cued digit-naming task involving both Chinese and English. Following the naming task, the participants performed the same antisaccade task again in the post-test. The results showed that the bilingual naming task enhanced response suppression, impeded the inhibitory control, and made no influence on the performance of task switching. Therefore, the authors suggest that moderate proficient bilinguals may rely heavily on response suppression when making speech production in two languages. As a consequence, the bilingual naming task improved their performance in the antisaccade task. Meanwhile, because of the limited cognitive resources and more involvement of response suppression, inhibitory control might have been allocated with less cognitive resources when moderate proficient bilinguals name digits using alternating languages, explaining the decreased performance in inhibitory control. Task switching involves a different mechanism from response suppression and inhibitory control and was not influenced by the bilingual context.

To reconcile discrepancies in previous studies, the present study explores the effect of language contexts on the neurocognitive mechanism of inhibitory control in a group of Cantonese–Mandarin–English trilinguals, who were highly proficient in Cantonese (L1)[1] and Mandarin (L2), and moderate proficient in English (L3). One possibility is that the discrepancies between Wu and Thierry (2013) and Liu et al. (2016) are not necessarily contradictory; they might arise as a result of differences in the participants' language background. The Welsh-English bilingual participants in Wu and Thierry (2013) were highly proficient in both languages; in contrast, the Chinese–English bilinguals in Liu et al. (2016) were intermediate learners of English. Bilinguals with high and low levels of L2 proficiency might adopt different processing strategies during speech production and, therefore, have incomparable implications for executive functions. In the same vein, age of L2 acquisition could also explain discrepancies between the two studies. Early and late bilinguals might engage different cognitive and neural mechanisms during language processing, so that the effect of language context on executive control might not be comparable between the two types of bilinguals. Finally, it is worth noting that unlike Chinese and English, Welsh and English are both alphabetical languages. Switching between two languages with more similarities in linguistic structures might engage different executive components as compared to switching between two languages that differ more radically.

To verify that language contexts may exert different effects on the inhibitory control of bilinguals with different language backgrounds, the present study examined trilingual speakers while they performed a flanker task (Luk et al., 2010) following picture naming in different dual-language contexts: the L1-L2 context, the L2-L3 context, and the L1-L3 context. Within-subject comparisons of their performance in the flanker task following the three contexts will provide a more confident answer to the modulation effect of language context on inhibitory control. It is our hypothesis that in the L1-L2 context (Cantonese–Mandarin), as in Wu and Thierry (2013), bilingual context would facilitate inhibitory control performance; in the L2-L3 (Mandarin–English) and the L1-L3 (Cantonese–English) contexts, as in Liu et al. (2016), the bilingual context would have no beneficial effect on inhibitory control.

The second goal of the current study is to examine how dual-language contexts modulate the functional brain network for inhibitory control, which is usually right-lateralized. For this purpose, effective connectivity analyses, following a recently developed procedure for valid group modeling, namely Group Interactive Multiple Model Estimation (GIMME, Gates and Molenaar, 2012) was performed to identify causal relationships between key brain regions that subserve inhibitory control in different dual-language contexts. If dual-language contexts do not influence inhibitory control process, participants should display a

---

[1]Although Mandarin and Cantonese are often referred to as two dialects of Chinese, they are mutually unintelligible in their oral forms (Tang and van Heuven, 2009), and significantly different in phonology, lexicon, and syntax (Cai et al., 2011). Therefore, in the literature of bilingualism, Mandarin and Cantonese represent two distinct languages, rather than two dialects (e.g., Cai et al., 2011; Tu et al., 2015; Liu et al., 2017).

typical flanker effect and comparable brain activation patterns as well as common functional brain network when performing the flanker task. If dual-language contexts do modulates inhibitory control, it is our hypothesis that the L1-L2 context would elicit a right-lateralized network for inhibitory control, while the L1-L3 and L2-L3 contexts might engage a less typical inhibitory control network, because of the more demanding task on linguistic processing and language control in the L2-L3 and L1-L3 contexts, relative to the L1-L2 context.

MATERIALS AND METHODS

Participants

Thirty students (10 males; age range 18–25) were recruited from the Guangdong University of Foreign Studies in Guangzhou, a city with a large Cantonese–Mandarin bilingual community. All participants were highly proficient early bilinguals of Cantonese (first language, L1) and Mandarin (second language, L2): They were raised up in a Cantonese family and have acquired Mandarin since early childhood. At the time of testing, participants use both languages on a regular basis.

All participants were late learners of English (third language, L3) in the mainstream classroom and had a moderate level of proficiency. They started to learn English at an average age of 7.4 (±1.82). According to their self-report, English and Mandarin were used as the main instruction languages in their English class (English usage: 52% ± 0.22; Mandarin usage: 40% ± 0.2; Cantonese usage: 7% ± 0.11), implying considerable experiences of switching and translation between English and Mandarin as a result of English learning.

As shown in **Table 1**, to assess the participants' linguistic knowledge and background variables in each of their three languages, we asked them to complete the following measures: (1) responses to the Language History Questionnaire (LHQ 2.0; Li et al., 2014) including the age of language acquisition (AoA), usage habits, switching frequency, and language abilities, (2) vocabulary knowledge in each language as examined through naming accuracy rates in a picture naming task (48 out of the 96 high-frequency non-living objects were selected as the stimuli from the battery of Snodgrass and Vanderwart (1980) and matched between languages, and (3) the Oxford Quick Placement Test (2001) as measurements of their English proficiency.

Based on language experience, usage habits, and language proficiency, the participants in the current study were characterized as (1) highly proficient in Cantonese and Mandarin, and with extensive experiences of switching between the two languages during conversations (i.e., in the L1-L2 context), and (2) moderately proficient in English but with more Mandarin–English switching experience (i.e., in the L2-L3 context) than Cantonese–English switching experience (i.e., in the L1-L3 context). All participants were right-handed as measured by the handedness inventory (Snyder and Harris, 1993). Written informed consent was obtained from all participants prior to the experiment. The Human Research Ethics Committee for Non-Clinical Faculties at the School of Psychology of South China Normal University approved this

study. All subjects gave written informed consent in accordance with the Declaration of Helsinki.

Materials and Procedure
Cognitive Assessments

Before the fMRI session, all participants received a battery of behavioral tests that are designed to measure their non-verbal intelligence (the Raven's Standard Progressive Matrices; Raven et al., 1988) and working memory (the odd-even sequencing task, an adaptation of number-sequencing subtest form the WAIS-III; Wechsler, 1997) as shown in **Table 1**.

fMRI Procedure

Participants completed six event-related fMRI runs, each lasting 6 min and 36 s. As shown in **Figure 1B**, every picture-naming run was presented prior to a flanker run. The order of the three dual-language contexts (i.e., L1-L2, L2-L3, and L1-L3) was counterbalanced between participants in a Latin square design.

The Picture Naming Task: As shown in **Figure 1A**, picture-naming tasks in the three dual-language contexts followed

TABLE 1 | Demographic variables, measures of cognitive skills, and language background information of the Cantonese–Mandarin–English trilingual participants.

	Cantonese–Mandarin–English Trilinguals (n = 30)
Age	21.64 ± 1.34
Handedness	42.77 ± 2.31
Non-verbal IQ (%)	74.17% ± 24.46
Working Memory (max:21)	20.03 ± 0.96
Processing Speed (ms)	1307.03 ± 180.21
Language measures	
Oxford Quick Placement Test (max:60)	45.4 ± 5.88
L1 picture naming (ACC)	99% ± 0.01
L2 picture naming (ACC)	100% ± 0.01
L3 picture naming (ACC)	95% ± 0.07
Language History Questionnaire (self-report)	
Reading in L1 (max:7)	6.7 ± 0.75
Writing in L1 (max:7)	6.17 ± 1.09
Speaking in L1 (max:7)	6.87 ± 0.43
Listening in L1 (max:7)	6.93 ± 0.25
Years of L1 learning	20.07 ± 1.41
Reading in L2 (max:7)	6.73 ± 0.58
Writing in L2 (max:7)	6.77 ± 0.57
Speaking in L2 (max:7)	6.57 ± 0.73
Listening in L2 (max:7)	6.87 ± 0.35
Years of L2 learning	17.6 ± 1.87
Reading in L3 (max:7)	5.7 ± 0.88
Writing in L3 (max:7)	5.13 ± 0.73
Speaking in L3 (max:7)	4.9 ± 0.96
Listening in L3 (max:7)	4.93 ± 0.83
Years of L3 learning	13.83 ± 2.15

L1, the first language (Cantonese); L2, the second language (Mandarin); L3, the third language (English). RT: reaction time; ACC: accuracy rate.
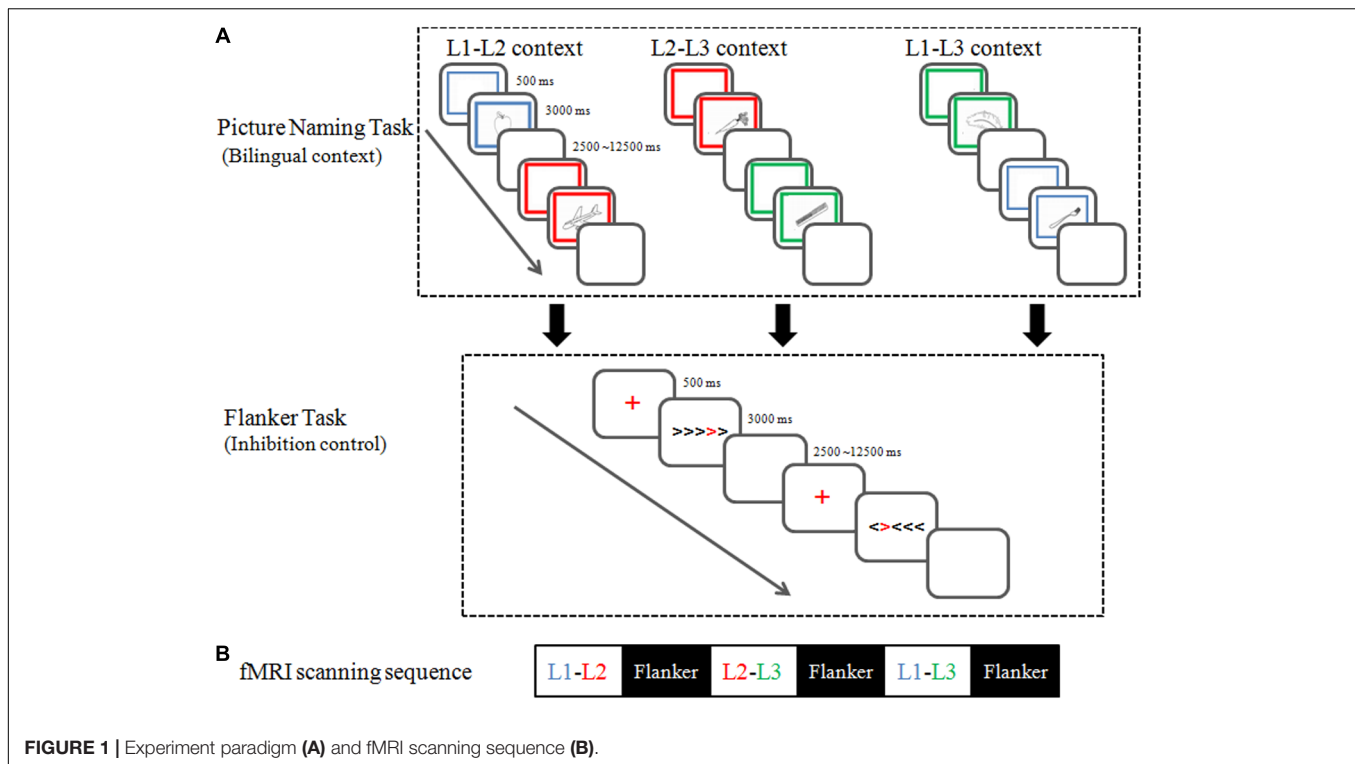
FIGURE 1 | Experiment paradigm **(A)** and fMRI scanning sequence **(B)**.

the same paradigm, each involving two different languages. During the picture-naming task, participants named pictures in two languages alternatingly, with 24 pictures per language. The stimuli were randomly selected from 96 black and white drawings for concrete non-living objects (e.g., piano) in the UCSD International Picture Naming Project (IPNP) picture database[2] (Bates et al., 2003). All stimuli corresponded to high frequency words in both Chinese (Liu et al., 2011) and English (Brysbaert and New, 2009), and were matched for word frequency (t_{95} = 0.4, p = 0.69) between the two languages. To ensure the familiarity of the object names, we asked an independent group of 35 individuals (age range: 18–21) from the same population to rate the familiarity of the object names on a 5-point Likert scale (1 = very infrequent; 5 = very frequent). There were no significant differences between the levels of familiarities of object names in the three languages [L1 = 4.0 ± 0.62, L2 = 4.13 ± 0.71, L3 = 3.94 ± 0.74; $F_{(2,54)}$ = 2.65, ps > 0.05].

In each trial of the picture naming task, a frame was presented for 500 ms and then a picture of an object appeared in the center of the frame for 3000 ms, followed by a blank screen of 500 ms. The color of the frame (blue, red, and green) served as the naming cue (blue for Cantonese, red for Mandarin, and green for English). Participants were instructed to covertly name the picture in the target language within 3000 ms. The 48 picture-naming trials were presented in a pseudo-random order with a jittered inter-stimulus interval (min = 2000 ms, max = 12000 ms) optimized with OptSeq2 (Dale, 1999).[3] During the inter-trial

interval, a central fixation cross was presented. Due to equipment limitation and to minimize head movements during naming, naming responses were collected outside of the MRI scanner. As an orientation procedure, we informed participants that they would later complete a behavioral test related to the naming task inside the scanner. We collected behavioral data with the same task outside of the scanner 2 weeks after the fMRI sessions (e.g., Zou et al., 2012; Li et al., 2013).

The Flanker Task: Immediately following each picture-naming task, participants were scanned in a flanker task session, to examine the influence of language context on their inhibitory control process. During the flanker task (Luk et al., 2010), participants responded to the direction of a red chevron (i.e., the target) surrounded by other black chevrons (i.e., the flankers). As shown in **Figure 1A**, in congruent trials, flanker chevrons point in the same direction as the target, whereas in incongruent trials, flankers pointed in the opposite direction to the target. Twenty-four congruent trials and 24 incongruent trials were randomly presented during each flanker task scanning session, with jittered inter-stimulus intervals (min = 2000 ms, max = 12000 ms). Each trial began with the presentation of a red fixation for 500 ms, followed by the stimulus for 3000 ms, and then a blank buffer of 500 ms.

MRI Acquisition

MRI images were acquired on a 3-T scanner (Siemens Trio Tim) equipped with a 12-channel phased-array head coil at the South China Normal University, using a T2*-weighted gradient-echo EPI sequence (TE = 30 ms; TR = 2s; flip angle = 90°; slices = 32; matrix size = 64 × 64; FoV = 192 mm × 192 mm;

thickness = 4 mm). Participants lay supine in the scanner and viewed the visual stimuli via a back-projection mirror, while their heads were immobilized with cushions. For each run, the functional scanning was always preceded by 6 s of dummy scans (fixation) to ensure tissue steady-state magnetization. High-resolution (1 mm × 1 mm × 1 mm) anatomical images were acquired using a T1-weighted, 3D inversion-recovery gradient-echo (MP-RAGE) sequence.

Data Analyses

Whole Brain Activations

The fMRI data were preprocessed using the Statistical Parametric Mapping (SPM) software running under MATLAB (SPM12; Wellcome Department of Imaging Neuroscience, University College London).[4] All three flanker runs followed the same data processing procedure. The first three scans (dummy scans) of the 198 volumes collected were discarded to allow for T1 equilibration. The remaining 195 volumes were then realigned to the first volume for head-motion correction, co-registered to the individual anatomical images and then to the EPI template in SPM12 based on the Montreal Neurological Institute (MNI) stereotactic space, and resampled into 3 mm × 3 mm × 3 mm cubic voxels. The head motion and rotation of all participants were less than 3 mm of displacement or 3° of rotation.

For each participant, functional images collected from each flanker run were grouped into congruent and incongruent conditions. Individual brain activations corresponding to congruent or incongruent conditions (in contrast to fixation) were analyzed using general linear model (GLM) and were entered into the second level of group analysis to show the neural correlates underlying inhibitory control.

AlphaSim program in the REST (Song et al., 2011) software was used to correct for multiple comparisons in SPM (10,000 interactions). All the brain activations reported below survived an FWE-corrected cluster-level threshold of $p < 0.05$ (single voxel: $p < 0.001$, number of voxels > 12) (Woo et al., 2014) and were in the MNI coordinate space.

ROI Selection and Analysis

Based on previous fMRI literature on language control and inhibitory control (Garavan et al., 1999; Luk et al., 2012; Grant et al., 2015; Abutalebi and Green, 2016), we selected 12 regions of interest (ROIs) to compose a cortical-subcortical-cerebellar network, which includes the right middle frontal gyrus (MFG) (33, 36, 21), right inferior frontal gyrus (IFG) (48, 9, 21), bilateral insula (INS) (±36, 3, −3), bilateral supramarginal gyri (SMG) (±18, 0, 21), bilateral caudate nuclei (CN) (±18, −21, 24), bilateral thalamus (THA) (±21, −30, 3), and bilateral cerebellum (CERE) (±15, −69, 42). Averaged time course data of all the voxels within a sphere (6 mm radius) in each ROI were extracted using the DPBABI software (Yan et al., 2016)[5].

To identify activation changes in those regions between congruent and incongruent conditions following the three different dual-language contexts, the present study sorted time

series of the 12 ROIs by experimental conditions (e.g., congruent and incongruent). The averaged time course signals across all trials of the congruent/incongruent conditions were converted to percentage signal changes (PSC) using the formula (signal-baseline)/baseline × 100 for each time point, and the baseline constant was the mean signal of the fixation baseline (e.g., Li et al., 2013). The averaged PSC values for each condition in every context were considered as representative activation level of each ROI for every participant.

Effective Connectivity

To examine the influence of dual-language context on functional brain connectivity of inhibitory control process in the trilinguals, we made use of recent advances in connectivity modeling (extended unified structural equation modeling, euSEM) (Gates et al., 2011; Hillary et al., 2011, 2014; Gates and Molenaar, 2012; Yang et al., 2015) and a recently developed Group Iterative Multiple Model Estimation (GIMME), an automatic and freely distributed MATLAB-based program[6].

The euSEM approach has provided a flexible and efficient method for analyzing the causal interactions of brain regions for cognitive functions, as has previously been applied in Grant et al. (2015) and Yang et al. (2015). The procedure for using the euSEM in the current study is consistent with Yang et al. (2015), but with two experimental conditions in every language context, namely congruent and incongruent conditions of the flanker task. As with other SEM-based approaches, GIMME works from individual-level correlation matrices. The covariance matrices used for the euSEM analysis include the ROI time series at time t (contemporaneous series, where each "t" is a single brain volume or TR) and the same ROI time series at the next time t + 1 (lagged series). For the euSEM analysis, the covariance matrices also include two time series of the effects of the task inputs (congruent and incongruent) for both time t and t + 1, convolved with a canonical hemodynamic response function. In addition, the bilinear series can be used to measure the influence of task inputs on the relationship between ROIs by examining time series of each ROI at each time t multiplied by the convolved task input series at time t. The model selections at the group and individual levels are implemented in the following steps. First, Lagrange Multiplier equivalents (i.e., modification indices; Sörbom, 1989) are used to identify which effects (including connections among ROIs, the direct and bilinear effects), if freed, optimally improve model fit *across all individuals*. The probability of detecting an effect across all individuals was set at 75%; selection of this criterion was informed by empirical and simulated studies on the likelihood of detecting a true effect should it exist in a given sample (e.g., Hillary et al., 2011, 2014; Gates and Molenaar, 2012; Yang and Li, 2012; Yang et al., 2015). The program iterates until the 75% criterion is met. Second, the model is pruned by eliminating connections that are no longer significant for 75% of the group after other connections are freed. Third, individual-level models are estimated in a semi-confirmatory manner. All connections freed in the group model (described in the two steps above) are

---

[4]http://www.fil.ion.ucl.ac.uk/spm

[5]http://rfmri.org/dpabi

[6]www.mathworks.com

freed at the individual level. The automatic search procedure within LISREL (Cziráky, 2004) then iteratively frees connections that optimally improve model fit, according to the Lagrange Multiplier equivalents (Gates et al., 2010). Finally, the model is pruned by eliminating individual-level connections that become non-significant after other individual-level connections are freed, and a confirmatory model is fitted. Model fit parameters, that were found to demonstrate reliability in simulation studies (e.g., Gates et al., 2010) and fMRI studies (e.g., Hillary et al., 2014), were chosen *a priori* so that two of the following four criteria were satisfied in the final model: confirmatory fit index (CFI) $\geq$ 0.9; non-normed fit index (NNFI) $\geq$ 0.9.

RESULTS

Behavioral Results

As shown in **Figure 2**, participants performed more quickly in the congruent condition as compared to the incongruent condition, displaying a typical flanker interference effect (Eriksen and Eriksen, 1974) in all three dual-language contexts ($ps < 0.001$). However, there were no significant differences in response times between the congruent or incongruent trials following the three dual-language contexts ($ps > 0.05$).
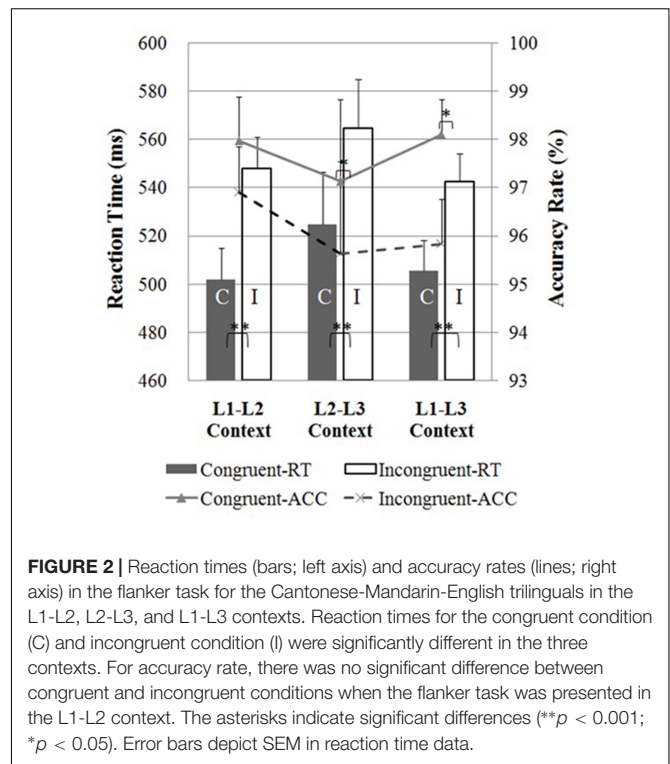
For accuracy rates, the flanker interference effect (higher accuracy rates in the congruent as compared to the incongruent trials) was found following the L2-L3 context (two-sample t-test, $t_{29} = 2.27$, $p < 0.05$) and the L1-L3 context ($t_{29} = 2.32$, $p < 0.05$), but not the L1-L2 context (two-sample t-test, $t_{29} = 1.14$, $p = 0.26$), suggesting a facilitation effect of the L1-L2 context on inhibitory control process.

fMRI Results

Whole-Brain Analysis

As shown in **Table 2**, in the L1-L2 context, incongruent trials elicited additional brain activations in the left inferior parietal lobe. In the L2-L3 context, incongruent conditions involved extra neural responses in the right inferior frontal gyrus and left supramarginal gyrus. In the L1-L3 context, both congruent and incongruent trials evoked neural activities in the right prefrontal cortex, right insula, and subcortical areas. See Supplementary Figure S1 for brain activations during the naming tasks.

No significant differences of neural activations were found between congruent and incongruent conditions in the L1-L2 and L2-L3 contexts. In the L1-L3 context, incongruent condition, as compared to the congruent condition, showed more neural responses in the bilateral inferior occipital gyri (Brodmann area 19, or BA 19), right middle occipital gyrus and bilateral middle temporal gyri (MTG). It is well-known that the medial temporal lobe (MTL) is the hub for declarative memory and keeps semantic representation (Squire et al., 2004). MTG might be a multimodal semantic processing hub, storing long-term conceptual knowledge, processing lexico-semantic information, and fulfilling semantic integration, especially in the L2 lexical processing (Rodríguez-Fornells et al., 2009). The stronger activation of the bilateral MTG in incongruent trials following L1-L3 context might imply competition of cognitive resources

FIGURE 2 | Reaction times (bars; left axis) and accuracy rates (lines; right axis) in the flanker task for the Cantonese-Mandarin-English trilinguals in the L1-L2, L2-L3, and L1-L3 contexts. Reaction times for the congruent condition (C) and incongruent condition (I) were significantly different in the three contexts. For accuracy rate, there was no significant difference between congruent and incongruent conditions when the flanker task was presented in the L1-L2 context. The asterisks indicate significant differences (**$p < 0.001$; *$p < 0.05$). Error bars depict SEM in reaction time data.

between inhibitory control task and the demanding semantic processing in L1-L3 context. As illustrated in **Figure 3A**, dual-language contexts (i.e., the L2-L3 and L1-L3 contexts) involving a moderate proficient language (L3) displayed increased brain activity in the right prefrontal cortex, bilateral insula and inferior parietal lobules, as well as subcortical areas, particularly the bilateral caudate and putamen, as compared to the L1-L2 context.

ROI Analyses

The following 12 ROIs were chosen based on the extant imaging literature of inhibitory control and language control (see ROI Selection in the Section of Materials and Methods): The right MFG, right IFG, bilateral INS, SMG, CN, Thalamus (THA), as well as bilateral cerebellum (CERE). Our analyses of the percent BOLD signal changes in those ROIs found (1) significant flanker effects in all the three dual-language contexts in the right cerebellum [right CERE, $F_{(1,29)} = 15.56$, $p < 0.001$]; (2) brain activations in the right IFG and left cerebellum were associated with the flanker effect in the L2-L3 context [right IFG, $F_{(1,29)} = 5.65$, $p = 0.024$; left CERE, $F_{(1,29)} = 6.19$, $p = 0.019$]; (3) neural responses in the right cerebellum, right MFG, right IFG, and left INS were associated with the flanker effect only in the L1-L3 context [MFG, $F_{(1,29)} = 7.09$, $p = 0.013$; IFG, $F_{(1,29)} = 58.75$, $p = 0.006$; INS, $F_{(1,29)} = 4.37$, $p = 0.045$] (**Figure 3B**).

Connectivity Analysis

An extended unified Structural Equation Model (euSEM) analysis was conducted on the fMRI data of the flanker

TABLE 2 | Whole brain activations associated with the flanker conditions for the Cantonese–Mandarin–English trilinguals in the L1-L2, L2-L3 and L1-L3 contexts.

	Congruent > fixation						Incongruent > fixation					
	Peak MNI Coordinate				Voxel		Peak MNI Coordinate				Voxel	
	BA	x	y	z	size	t	BA	x	y	z	size	t
L1-L2 context												
Insula	–	−42	0	9	32	5.1						
Postcentral gyrus	1	−63	−18	30	131	6.3	1	−54	−21	27	30	5.2
	1	60	−15	24			1	63	−15	21	18	5.3
Supramarginal gyrus							40	−60	−21	42	12	4.9
Fusiform gyrus	37	27	−48	−18	1857	8.9						
Lingual gyrus							18	15	−87	−12	1437	9.1
Thalamus	–	−21	−30	3	15	7.1	–	−21	−30	3	15	6.7
L2-L3 context												
Inferior frontal gyrus							44	48	9	21	15	5.5
Insula	–	−36	3	−3	16	4.7						
	–	36	6	3	15	5.0	–	45	9	−3	53	5.9
Rolandic operculum							–	−39	−3	15	25	5.5
Postcentral gyrus	1	−60	−18	24	52	5.1						
Supramarginal gyrus	40	57	−21	21	42	4.8	40	−63	−21	21	52	5.8
							40	57	−21	21	180	5.6
Lingual gyrus							18	15	−78	−12	2061	7.7
Fusiform gyrus	37	21	−48	−15	2037	9.6						
Thalamus	–	24	−30	0	15	6.7	–	24	−30	0	19	6.5
L1-L3 context												
Middle frontal gyrus	9	33	36	21	17	5.0	9	33	36	21	23	5.3
Inferior frontal gyrus	44	60	15	15	29	4.9	44	51	15	0	12	4.3
							45	45	39	3	26	5.2
Insula	–	−27	−3	−9	13	4.9						
	–	39	3	0	14	5.5	–	36	0	12	18	4.6
Rolandic operculum							–	−48	−9	12	29	5.4
Postcentral gyrus	1	−54	−21	39	119	5.7						
	1	54	−24	54	21	5.2						
Heschl gyrus	13	−48	−12	9	33	5.1						
Superior parietal lobule							7	21	−69	57	24	5.2
Inferior parietal lobule							40	−48	−39	57	234	6.2
Precuneus							7	12	−48	75	16	4.2
Supramarginal gyrus							40	60	−18	27	212	7.0
							40	30	−39	42	21	4.4
Lingual gyrus	18	9	−78	−9	1722	8.1	18	9	−78	−9	2371	8.4
Caudate	–	−18	−15	24	25	5.6	–	−18	−21	24	22	5.0
	–	18	−21	21	25	5.6	–	18	−15	24	14	5.1
Putamen	–	27	−3	−12	18	5.3		−30	−12	−9	24	5.3
							–	33	−15	−6	16	6.3
Hippocampus							–	−24	−30	−3	15	5.3
							–	18	−30	−3	32	6.4
Cerebellum	–	15	−72	−42	12	4.7	–	−3	−69	−36	14	4.9
							–	15	−69	−42	14	4.4

tasks following the L1-L2, L2-L3 and L1-L3 contexts. All group maps (**Figure 4**) had an excellent fit to the data for roughly 97–100% of the participants, depending on the measure (Brown, 2006). Specifically, in the L1-L2 context, the Comparative Fit Index (CFI) evaluated the model fit as excellent for 100% of the participants' data, while Standardized ROOT Mean Square Residual (SRMR) and the Root Mean Square Error of Approximation (RMSEA) showed excellent fit for 97% of the data. In L2-L3 context, CFI, SRMR and the RMSEA results indicated an excellent fit for 100% of the data, as is the same in the L1-L3 context.
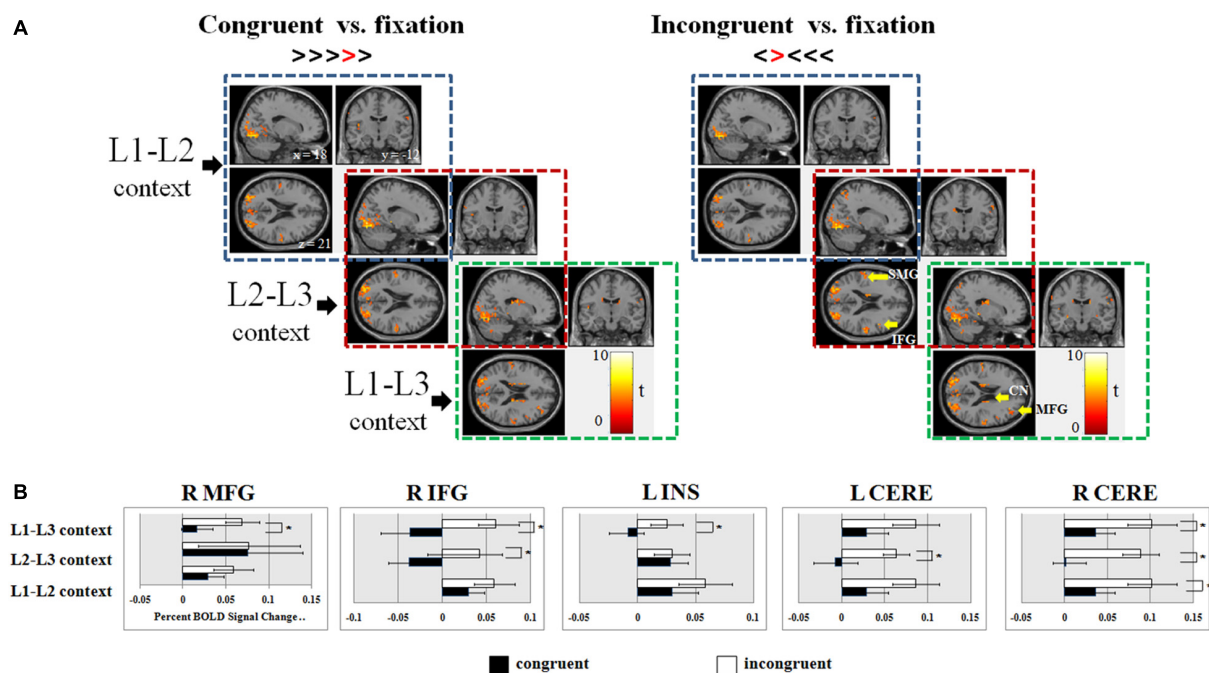
FIGURE 3 | (A) Overall brain activations associated with congruent and incongruent trials of the flanker task as presented in the L1-L2, L2-L3, and L1-L3 contexts; **(B)** Regions of interests that were sensitive to the flanker effect in different dual-language contexts. MFG, middle frontal gyrus; IFG, inferior frontal gyrus; INS, insula; CN, caudate nucleus; THA, thalamus; SMG, supramarginal gyrus; CERE, cerebellum; L, left hemisphere; R, right hemisphere; *$p < 0.05$.
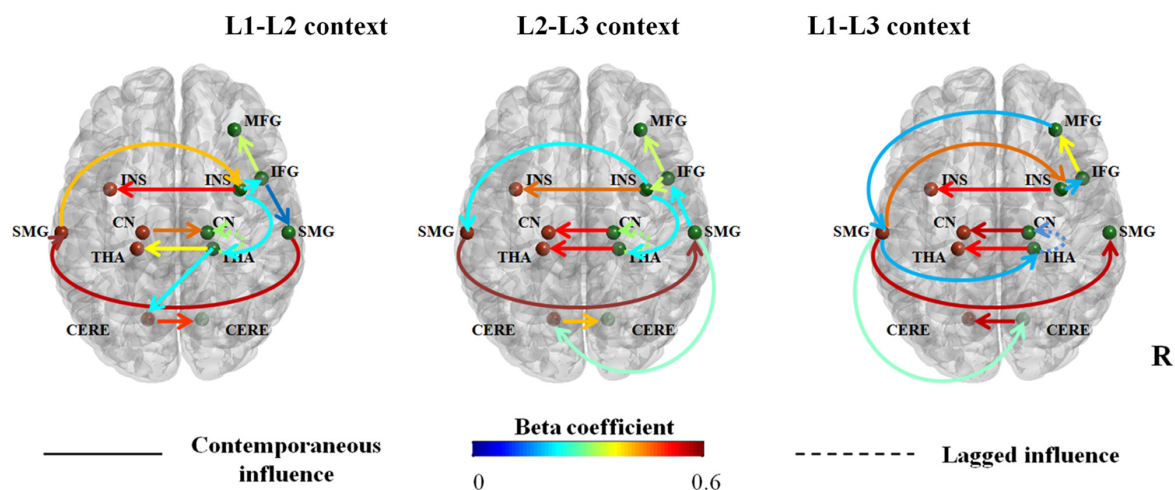

FIGURE 4 | Connectivity maps for inhibitory control process in three dual-language contexts (L1-L2; L2-L3; L1-L3). Significant contemporaneous and lagged relationships at group-level were shown, and auto-regressive lagged connections are omitted for clarity. Connection strength is denoted by beta coefficients, reflected here as the color the line. The solid line indicates contemporaneous relationship, namely area X at time T influences brain activation of area Y at time T; dotted line illustrates lagged relationship, namely area X at time T influences brain activation of area Y at time T + 1. Nodes (ROIs) are MFG, middle frontal gyrus; IFG, inferior frontal gyrus; INS, insula; CN, caudate nucleus; THA, thalamus; SMG, supramarginal gyrus; CERE, cerebellum; R, right hemisphere.

As shown in **Figure 4**, the cortical-subcortical-cerebellar network for inhibitory control following the three dual-language contexts shared common connections: The right INS and thalamus strongly influence their left homologous areas, implying the right-dominant network for inhibitory control; the right IFG feeds to the right MFG, suggesting that the right IFG is highly engaged in bottom–up process of inhibitory control; the right thalamus influences brain activations in the right caudate in a lagged relationship, implying the key role of right thalamus in the communications between cortical and subcortical areas. In all the three contexts, right INS influences right IFG and left SMG. Not surprisingly, the left and right

SMG are connected to each other as for the bilateral caudate nuclei.

Obviously, in all three dual-language contexts, inhibitory control relied on collaborations between a frontal-parietal network, a cortical-thalamic-striato pathway and bilateral cerebellum. However, dual-language contexts influence inhibitory control network in the trilinguals. To be specific, inhibitory control in the L1-L2 context recruits an efficient and right lateralized network: The right INS as the hub of the network feeds to the right IFG, modulating brain activation in the right MFG indirectly and activating the right SMG; the right THA, as a mediator of frontal-thalamic-striato pathway, receiving positive information from the right INS, feeds to the right CN; right THA is also a key relay station for cortico-cerebellar pathway, forwarding information from the right INS to the left cerebellum, which sends strong and positive influence to the right cerebellum; right SMG forward information from the right IFG to the left SMG, which passes information to the right insula, completing the frontal-parietal sub-network.

When the dual-language context involves a moderate proficient language (L3), the functional brain network for inhibitory control in the same group of participants changed immediately. Specifically, in the L2-L3 context, the frontal-parietal sub-network runs in the reverse pipeline: The right IFG actively influences the right INS, which as a hub sends direct and positive information to the left SMG, feeding directly to the right SMG; the right IFG receives feedback from the right SMG and connects to the right MFG. As in the L1-L2 context, right INS influences bilateral caudate via THA in a lagged relationship. The left cerebellum receives inputs directly from the right SMG and feeds to the right cerebellum. Interestingly, the right CN activates left CN in the L2-L3 context, as is the same in the condition of the L1-L3 context.

L1-L3 context involves a moderate proficient language and engages less switching experience between the two languages. Compared with the L1-L2 and L2-L3 contexts, inhibition control in the L1-L3 context relies on a more left-lateralized and less integrated network, compared with the networks in the other two dual-language contexts. In L1-L3 context, the left SMG works as the hub of the inhibitory control network: The right MFG takes the lead and sends orders directly to the left SMG, which relays to the right INS, right SMG, right THA, and right cerebellum. The right THA, as in the other two conditions, forward positive information to the right CN in a lagged relationship, thus completing the cortico-thalamic-striato pathway. For the cerebellar components, the right cerebellum receives weak influence from SMG on the contralateral side and feeds to the left cerebellum.

Taken together, the cortical-subcortical-cerebellar network for inhibitory control involves a frontal-parietal sub-network, cortico-thalamic-striato pathway, and bilateral cerebellum. However, this neural pattern can be modulated by language contexts on a short timescale. In dual-language contexts with intensive code switching between two proficient languages (i.e., the L1-L2 context), the inhibitory control process seems to be facilitated and rely on a right-lateralized control network. When the dual-language context involves a less proficient language, especially when the two languages are radically different from each other, the inhibitory control process relies on a more left-lateralized and less integrated neural network: In the L2-L3 context, the right IFG feeds to right INS and receives feedback from the right SMG; in the L1-L3 context, the right MFG is highly engaged and the whole network relies on the left SMG to modulate brain activations in the cortical, subcortical, and cerebellar areas.

DISCUSSION

The current study examines the dynamic influences of language contexts on inhibitory control in trilinguals. We explored the neural correlates and functional brain networks activated while Cantonese–Mandarin–English trilingual speakers performed the flanker tasks in three dual-language contexts. As bilingual language processing engages inhibition of the non-target language (Dijkstra and Van Heuven, 1998; Green, 1998; Linck et al., 2008; Green and Abutalebi, 2013; Ventura-Campos et al., 2013; Abutalebi and Green, 2016) and acquiring a second language facilitates the development of cognitive control (Linck et al., 2009; Hosoda et al., 2013; Grant et al., 2015), we expected a facilitatory effect of dual-language contexts on participants' cognitive control performance (e.g., Wu and Thierry, 2013). Our participants were highly proficient in L1 and L2, but moderately proficient in L3. They frequently switch between L1 and L2 in everyday life, but not between L1 and L3 or L2 and L3. We, therefore, expected to observe significant facilitatory effect of the L1-L2 context as compared to the L2-L3 and L1-L3 contexts on inhibitory control.

The results showed the classic flanker effect in the L2-L3 and L1-L3 contexts, but not in the L1-L2 context. Consistent with results in a previous study (Wu and Thierry, 2013), the effect of contextual priming was observed in accuracy rates but not reaction times, suggesting that independent cognitive mechanisms might account for flanker effects in the two types of measurements. Previous studies (e.g., Luk et al., 2010) using a similar flanker task showed that in the incongruent condition, bilinguals activated a widespread set of brain regions, including the fusiform gyri, inferior frontal gyri, supplementary motor area, inferior parietal regions, and subcortical areas. In the present study, those brain regions for inhibitory control failed to show significant activations in incongruent condition (when compared to the congruent condition) in the L1-L2 context. Moreover, none of the ROIs showed significant interference effects in the L1-L2 context in terms of their BOLD signal change (**Figure 3B**). These findings suggest that the neural efficiency of the inhibitory control network was enhanced in the L1-L2 context, reducing the classic flanker effect in both the behavioral and the neural anatomical level (e.g., Jäncke et al., 2000; Stevens et al., 2007). Effective connectivity analysis shows the following pattern of results: The right insula functions as the hub of the frontal-parietal network, feeding to the right IFG, which mediates the right MFG and bilateral SMG; the right MFG and bilateral SMG then send information to the right THA, which positively modulates brain activations of bilateral caudate nuclei in a lagged relationship and

directly influences the cerebellar pathway. The important role of the right insula in inhibitory control has been well documented. In an event-related fMRI study, Garavan et al. (1999) showed that performing a response inhibition task activated the right hemisphere, including the right MFG, IFG, insula, and inferior parietal lobule. A more recent study dissociated the functional role of the right IFG and insula in inhibitory control and suggests that the right insula is particularly important for detecting behaviorally salient events, while the right IFG is more involved in implementing inhibitory control (Cai et al., 2014). Meanwhile, it is interesting to note that the right frontal-insular cortex has been implicated in switching between central-executive and default-mode networks (Sridharan et al., 2008).

In the L2-L3 context, behavioral results showed significant flanker effects in both reaction times and accuracy rates (see **Figure 2**). Analysis of the neuroimaging data showed that the flanker effect was associated with brain activations in the right inferior frontal gyrus, bilateral insula, left rolandic operculum, bilateral supramarginal gyrus, and right thalamus. This pattern of activations is highly consistent with neural mechanisms underlying typical flanker effects (e.g., Luk et al., 2010). ROI analyses showed that the flanker effect was associated with activations in the right IFG and left cerebellum, where no such significant activations were found in the L1-L2 context, suggesting a priming effect on the inhibitory control when the dual-language context involves two highly proficient languages. Connectivity map shows a similar frontal-parietal network, but in the reverse relationship. To be specific, in the L1-L2 context, the right IFG sends feedback to the right MFG and completes the frontal-parietal circuitry via the right SMG and left SMG; in the L2-L3 context, flanker task involves more engagement of the IFG: The IFG passes positive influences to the right INS, forwarding information to the left SMG and the right THA; the right SMG receives signals from the left SMG and sends feedback to the IFG, completing the frontal-parietal loop.

In the L1-L3 context, the flanker effect is associated with a different neural network, in which the left SMG is the hub. As illustrated in **Figure 4**, the right MFG influences the left SMG, which influences the right INS, the right IFG, and finally the right MFG, completing a frontal-insula-parietal network without the right SMG; meanwhile, the left SMG directly modulates brain activations in the right THA, which communicates with the right CN in a lagged relationship as in the other two contexts; furthermore, the left SMG feeds to the right cerebellum, which connects to the left cerebellum. This distinction of the neural network in the L1-L3 context, as compared to the L1-L2 and the L2-L3 contexts, is further supplemented by increased activations in the right MFG, right IFG, bilateral basal ganglia and cerebellum (see **Table 2**). ROI analyses also showed that activations in the right MFG and right cerebellum were associated with the flanker effect.

Research in neuropsychology and cognitive neuroscience has established the role of the left supramarginal gyrus in the inferior parietal lobule in second language acquisition. Neuroimaging data showed that early bilingualism is associated with increased gray matter density in the left inferior parietal

lobe (Mechelli et al., 2004). In addition, researchers have found that the lateral inferior parietal cortex contributes to attentional focalization and target detection in both auditory and visual modalities, indicating its involvement in domain-general attentional processes (e.g., Green et al., 2006; Shomstein and Yantis, 2006). As summarized by Della Rosa et al. (2013), second language acquisition might tune this attentional control area into a "multilingual talent area" as phonological storage and attentional control functions were also subserved by this left inferior parietal lobe. The right SMG, as suggested by Abutalebi and Green (2008), was particularly involved in language selection in conversations that involve multiple languages. As illustrated in **Figure 4**, in the L1-L2 context, the right SMG is influenced by the right IFG in a top-down control process, while in the L2-L3 context, it sends feedback to the right IFG, forming a bottom–up stream.

It is worth noticing that in the L1-L2 context, consistent with the adaptive control hypothesis (Abutalebi and Green, 2016), the right IFG feeds to the right insula, which influences the right thalamus, thus modulating subcortical areas such as the caudate and connecting the cerebellum. The thalamus has been reported to directly connect to the regions of basal ganglia (Smith et al., 2011) and has reciprocal structure connections with the cerebellum as a relay station (e.g., Glickstein and Doron, 2008). The left caudate and putamen might be more involved in verb interference effects (Abutalebi and Green, 2008; Ali et al., 2010), while the right homologous areas play a more important role in inhibitory control. Based on our results of trilinguals, dual-language contexts modulate the involvement of inhibitory control areas and their interactions.

CONCLUSION

The finding that dual-language contexts lead to functional reorganizations of the inhibitory control network not only reconciles discrepancies in previous studies (e.g., Wu and Thierry, 2013; Liu et al., 2016), but also provides a novel perspective for investigating the interplay between language control and non-linguistic cognitive processes. To fully understand the nature of the neural mechanisms subserving non-linguistic skills (e.g., executive functions), researchers have to consider the influences of processing contexts. Results of the current study provides empirical evidence in favor of the adaptive control hypothesis (Green and Abutalebi, 2013), which suggests that interactional contexts (e.g., single-language, dual-language, or frequent-switching) modulate language control processes by adaptive changes in the neural regions and circuits associated with specific control processes. Critically, our results showed that the cognitive system and its underlying neural network are highly plastic, allowing quick development of functional reconfigurations. Short-term language engagement, in the form of contextual priming, can instantly rewire the related brain mechanisms. This finding sheds new light on therapy training programs for individuals with minor cognitive impairment (MCI). Whether or not L2 proficiency, age of acquisition, and cross-language similarities (e.g., alphabetical or non-alphabetical)

distinctly contribute to the modulation effects of bilingual contexts requires further exploration.

REFERENCES

Abutalebi, J., Della Rosa, P. A., Ding, G., Weekes, B., Costa, A., and Green, D. W. (2013a). Language proficiency modulates the engagement of cognitive control areas in multilinguals. *Cortex* 49, 905–911. doi: 10.1016/j.cortex.2012.08.018

Abutalebi, J., Della Rosa, P. A., Gonzaga, A. K., Keim, R., Costa, A., and Perani, D. (2013b). The role of the left putamen in multilingual language production. *Brain Lang.* 125, 307–315. doi: 10.1016/j.bandl.2012.03.009

Abutalebi, J., and Green, D. W. (2007). Bilingual language production: the neurocognition of language representation and control. *J. Neurolinguistics* 20, 242–275. doi: 10.1016/j.jneuroling.2006.10.003

Abutalebi, J., and Green, D. W. (2008). Control mechanisms in bilingual language production: neural evidence from language switching studies. *Lang. Cogn. Process.* 23, 557–582. doi: 10.1080/01690960801920602

Abutalebi, J., and Green, D. W. (2016). Neuroimaging of language control in bilinguals: neural adaptation and reserve. *Bilingualism* 19, 689–698. doi: 10.1017/S1366728916000225

Ali, N., Green, D. W., Kherif, F., Devlin, J. T., and Price, C. J. (2010). The role of the left head of caudate in suppressing irrelevant words. *J. Cogn. Neurosci.* 22, 2369–2386. doi: 10.1162/jocn.2009.21352

Bates, E., D'Amico, S., Jacobsen, T., Székely, A., Andonova, E., Devescovi, A., et al. (2003). Timed picture naming in seven languages. *Psychon. Bull. Rev.* 10, 344–380. doi: 10.3758/BF03196494

Bercury, K. K., and Macklin, W. B. (2015). Dynamics and mechanisms of CNS myelination. *Dev. Cell* 32, 447–458. doi: 10.1016/j.devcel.2015.01.016

Bialystok, E. (2007). Acquisition of literacy in bilingual children: a framework for research. *Lang. Learn.* 57, 45–77. doi: 10.1111/j.1467-9922.2007.00412.x

Bialystok, E., Craik, F., and Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 859–873. doi: 10.1037/0278-7393.34.4.859

Bialystok, E., Craik, F. I., Grady, C., Chau, W., Ishii, R., Gunji, A., et al. (2005). Effect of bilingualism on cognitive control in the Simon task: evidence from MEG. *Neuroimage* 24, 40–49. doi: 10.1016/j.neuroimage.2004.09.044

Bialystok, E., Craik, F. I., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19, 290–303. doi: 10.1037/0882-7974.19.2.290

Bialystok, E., Craik, F. I. M., and Ryan, J. (2006). Executive control in a modified antisaccade task: effects of aging and bilingualism. *J. Exp. Psychol.* 32, 1341–1354. doi: 10.1037/0278-7393.32.6.1341

Brown, T. A. (2006). *Confirmatory Factor Analysis for Applied Research.* New York, NY: The Guilford Press.

Brysbaert, M., and New, B. (2009). Moving beyond Kuèera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behav. Res. Methods* 41, 977–990. doi: 10.3758/BRM.41.4.977

Cai, W., Ryali, S., Chen, T., Li, C.-S. R., and Menon, V. (2014). Dissociable roles of right inferior frontal cortex and anterior Insula in inhibitory control: evidence from intrinsic and task-related functional parcellation, connectivity, and response profile analyses across multiple datasets. *J. Neurosci.* 34, 14652–14667. doi: 10.1523/JNEUROSCI.3048-14.2014

Cai, Z. G., Pickering, M. J., Yan, H., and Branigan, H. P. (2011). Lexical and syntactic representations in closely related languages: evidence from Cantonese-mandarin bilinguals. *J. Mem. Lang.* 65, 431–445. doi: 10.1016/j.jml.2011.05.003

Costa, A., Hernández, M., and Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: evidence from the ANT task. *Cognition* 106, 59–86. doi: 10.1016/j.cognition.2006.12.013

Cziráky, D. (2004). LISREL 8.54: a program for structural equation modelling with latent variables. *J. Appl. Econ.* 19, 135–141. doi: 10.1002/jae.767

Dale, A. M. (1999). Optimal experimental design for event-related fMRI. *Hum. Brain Mapp.* 8, 109–114. doi: 10.1002/(SICI)1097-0193(1999)8:2/3<109::AID-HBM7>3.0.CO;2-W

De Groot, A. M. B. (2012). "Bilingualism and cognition," in *The Encyclopedia of Applied Linguistics*, ed. C. A. Chapelle (Hoboken, NY: Blackwell Publishing Ltd.).

Delekate, A., Zagrebelsky, M., Kramer, S., Schwab, M. E., and Korte, M. (2011). NogoA restricts synaptic plasticity in the adult hippocampus on a fast time scale. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2569–2574. doi: 10.1073/pnas.1013322108

Della Rosa, P. A., Videsott, G., Borsa, V. M., Canini, M., Weekes, B. S., Franceschini, R., et al. (2013). A neural interactive location for multilingual talent. *Cortex* 49, 605–608. doi: 10.1016/j.cortex.2012.12.001

Dijkstra, T., and Van Heuven, W. J. B. (1998). "The BIA model and bilingual word recognition," in *Localist Connectionist Approaches to Human Cognition*, eds J. Grainger and A. M. Jacobs (London: Psychology Press), 189–225.

Dong, Y., and Zhong, F. (2017). Interpreting experience enhances early attentional processing, conflict monitoring and interference suppression along the time course of processing. *Neuropsychologia* 95, 193–203. doi: 10.1016/j.neuropsychologia.2016.12.007

Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept. Psychophys.* 16, 143–149. doi: 10.3758/BF03203267

Fields, R. D. (2005). Myelination: an overlooked mechanism of synaptic plasticity? *Neuroscientist?* 11, 528–531.

Garavan, H., Ross, T. J., and Stein, E. A. (1999). Right hemispheric dominance of inhibitory control: an event-related functional MRI study. *Proc. Natl. Acad. Sci. U.S.A.* 96, 8301–8306. doi: 10.1073/pnas.96.14.8301

Gates, K. M., Molenaar, P., Hillary, F. G., Ram, N., and Rovine, M. J. (2010). Automatic search for fMRI connectivity mapping: an alternative to Granger causality testing using formal equivalences among SEM path modeling, VAR, and unified SEM. *Neuroimage* 50, 1118–1125. doi: 10.1016/j.neuroimage.2009. 12.117

Gates, K. M., Molenaar, P. C., Hillary, F. G., and Slobounov, S. (2011). Extended unified SEM approach for modeling event-related fMRI data. *Neuroimage* 54, 1151–1158. doi: 10.1016/j.neuroimage.2010.08.051

Gates, K. M., and Molenaar, P. C. M. (2012). Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples. *Neuroimage* 63, 310–319. doi: 10.1016/j.neuroimage.2012. 06.026

Ghazi Saidi, L., Perlbarg, V., Marrelec, G., Pélégrini-Issac, M., Benali, H., and Ansaldo, A. I. (2013). Functional connectivity changes in second language vocabulary learning. *Brain Lang.* 124, 56–65. doi: 10.1016/j.bandl.2012.11.008

Glickstein, M., and Doron, K. (2008). Cerebellum: connections and functions. *Cerebellum* 7, 589–594. doi: 10.1007/s12311-008-0074-4

Grant, A. M., Fang, S.-Y., and Li, P. (2015). Second language lexical development and cognitive control: a longitudinal fMRI study. *Brain Lang.* 144, 35–47. doi: 10.1016/j.bandl.2015.03.010

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 behavioral 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

Green, D. W., Crinion, J., and Price, C. J. (2006). Convergence. Degeneracy, and control. *Lang. Learn.* 56, 99–125. doi: 10.1111/j.1467-9922.2006.00357.x

Hillary, F. G., Medaglia, J. D., Gates, K., Molenaar, P. C., Slocomb, J., Peechatka, A., et al. (2011). Examining working memory task acquisition in a disrupted neural network. *Brain* 134(Pt 5), 1555–1570. doi: 10.1093/brain/awr043

Hillary, F. G., Medaglia, J. D., Gates, K. M., Molenaar, P. C., and Good, D. C. (2014). Examining network dynamics after traumatic brain injury using the extended unified SEM approach. *Brain Imaging Behav.* 8, 435–445. doi: 10.1007/s11682-012-9205-0

Hosoda, C., Tanaka, K., Nariai, T., Honda, M., and Hanakawa, T. (2013). Dynamic neural network reorganization associated with second language vocabulary acquisition: a multimodal imaging study. *J. Neurosci.* 33, 13663–13672. doi: 10.1523/JNEUROSCI.0410-13.2013

Jäncke, L., Shah, N., and Jand Peters, M. (2000). Cortical activations in primary and secondary motor areas for complex bimanual movements in professional pianists. *Cogn. Brain Res.* 10, 177–183. doi: 10.1016/S0926-6410(00)00028-8

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

Li, P., Zhang, F., Tsai, E., and Puls, B. (2014). Language history questionnaire (LHQ 2.0): a new dynamic web-based research tool. *Biling. Lang. Cogn.* 17, 673–680. doi: 10.1017/S1366728913000606

Li, Y., Yang, J., Suzanne Scherf, K., and Li, P. (2013). Two faces, two languages: an fMRI study of bilingual picture naming. *Brain Lang.* 127, 452–462. doi: 10.1016/j.bandl.2013.09.005

Linck, J. A., Hoshino, N., and Kroll, J. F. (2008). Cross-language lexical processes and inhibitory control. *Ment. Lex.* 3, 349–374. doi: 10.1075/ml.3.3.06lin

Linck, J. A., Kroll, J. F., and Sunderman, G. (2009). Losing access to the native language while immersed in a second language: evidence for the role of inhibition in second-language learning. *Psychol. Sci.* 20, 1507–1515. doi: 10.1111/j.1467-9280.2009.02480.x

Liu, C., Lu, J., Sun, X., and Wang, R. (2016). Immediate effect of language switch on non-proficient bilinguals' cognitive control components. *Acta Psychol. Sin.* 48, 472–481. doi: 10.3724/SP.J.1041.2016.00472

Liu, X., Tu, L., Wang, J., Jiang, B., Gao, W., Pan, X., et al. (2017). Onset age of L2 acquisition influences language network in early and late Cantonese-Mandarin bilinguals. *Brain Lang.* 174, 16–28. doi: 10.1016/j.bandl.2017. 07.003

Liu, Y., Hao, M., Li, P., and Shu, H. (2011). Timed picture naming norms for Mandarin Chinese. *PLoS One* 6:e16505. doi: 10.1371/journal.pone.0016505

Luk, G., Anderson, J. A., Craik, F. I., Grady, C., and Bialystok, E. (2010). Distinct neural correlates for two types of inhibition in bilinguals: response inhibition versus interference suppression. *Brain Cogn.* 74, 347–357. doi: 10.1016/j.bandc. 2010.09.004

Luk, G., Green, D. W., Abutalebi, J., and Grady, C. (2012). Cognitive control for language switching in bilinguals: a quantitative meta-analysis of functional neuroimaging studies. *Lang. Cogn. Process.* 27, 1479–1488. doi: 10.1080/ 01690965.2011.613209

Martin-Rhee, M. M., and Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Biling. Lang. Cogn.* 11, 81–93. doi: 10.1017/S1366728907003227

Mechelli, A., Crinion, J. T., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R. S., et al. (2004). Neurolinguistics: structural plasticity in the bilingual brain. *Nature* 431, 757–757. doi: 10.1038/431757a

Morton, J. B., and Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Dev. Sci.* 10, 719–726. doi: 10.1111/j.1467-7687.2007. 00623.x

Oxford Quick Placement Test (2001). *Oxford Quick Placement Test.* Oxford: Oxford University Press.

Raven, J. C., Raven, J., and Court, J. H. (1988). *The Mill Hill Vocabulary Scale.* London: H. K. Lewis.

Rodríguez-Fornells, A., Cunillera, T., Mestres-Missé, A., and de Diego-Balaguer, R. (2009). Neurophysiological mechanisms involved in language learning in adults. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 3711–3735. doi: 10.1098/ rstb.2009.0130

Sabbagh, C., Resh, N., Mor, M., and Vanhuysse, P. (2006). Spheres of justice within schools: reflections and evidence on the distribution of educational goods. *Soc. Psychol. Educ.* 9, 97–118. doi: 10.1007/s11218-005-3319-9

Shomstein, S., and Yantis, S. (2006). Parietal cortex mediates voluntary control of spatial and nonspatial auditory attention. *J. Neurosci.* 26, 435–439. doi: 10.1523/ JNEUROSCI.4408-05.2006

Smith, Y., Surmeier, D. J., Redgrave, P., and Kimura, M. (2011). Thalamic contributions to basal ganglia-related behavioral switching and reinforcement. *J. Neurosci.* 31, 16102–16106. doi: 10.1523/JNEUROSCI.4634-11.2011

Snodgrass, J. G., and Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *J. Exp. Psychol. Hum. Learn. Mem.* 6, 174–215. doi: 10.1037/0278-7393.6.2.174

Snyder, P. J., and Harris, L. J. (1993). Handedness, sex, familial sinistrality effects on spatial tasks. *Cortex* 29, 115–134. doi: 10.1016/S0010-9452(13)80216-X

Song, X.-W., Dong, Z.-Y., Long, X.-Y., Li, S.-F., Zuo, X.-N., Zhu, C.-Z., et al. (2011). REST: a toolkit for resting-state functional magnetic resonance imaging data processing. *PLoS One* 6:e25031. doi: 10.1371/journal.pone.0025031

Sörbom, D. (1989). Model modification. *Psychometrika* 54, 371–384. doi: 10.1007/ BF02294623

Squire, L. R., Stark, C. E., and Clark, R. E. (2004). The medial temporal lobe. *Annu. Rev. Neurosci.* 27, 279–306. doi: 10.1146/annurev.neuro.27.070203.144130

Sridharan, D., Levitin, D. J., and Menon, V. (2008). A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proc. Natl. Acad. Sci. U.S.A.* 105, 12569–12574. doi: 10.1073/pnas. 0800005105

Stevens, M. C., Kiehl, K. A., Pearlson, G. D., and Calhoun, V. D. (2007). Functional neural networks underlying response inhibition in adolescents and adults. *Behav. Brain Res.* 181, 12–22. doi: 10.1016/j.bbr.2007.03.023

Tang, C., and van Heuven, V. J. (2009). Mutual intelligibility of Chinese dialects experimentally tested. *Lingua* 119, 709–732. doi: 10.1016/j.lingua.2008.10.001

Thierry, G., and Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign language comprehension. *Proc. Natl. Acad. Sci. U.S.A.* 104, 12530–12535. doi: 10.1073/pnas.0609927104

Tu, L., Wang, J., Abutalebi, J., Jiang, B., Pan, X., Li, M., et al. (2015). Language exposure induced neuroplasticity in the bilingual brain: a follow-up fMRI study. *Cortex* 64, 8–19. doi: 10.1016/j.cortex.2014.09.019

Ventura-Campos, N., Sanjuán, A., González, J., Palomar-García, M.-Á., Rodríguez-Pujadas, A., Sebastián-Gallés, N., et al. (2013). Spontaneous brain activity predicts learning ability of foreign sounds. *J. Neurosci.* 33, 9295–9305. doi: 10.1523/JNEUROSCI.4655-12.2013

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale-(WAIS-III)*. San Antonio, TX: Psychological Corporation.

Weissberger, G., Gollan, T., Bondi, M., Clark, L., and Wierenga, C. (2015). Language and task switching in the bilingual brain: bilinguals are staying, not switching, experts. *Neuropsychologia* 66, 193–203. doi: 10.1016/j.neuropsychologia.2014.10.037

Woo, C.-W., Krishnan, A., and Wager, T. D. (2014). Cluster-extent based thresholding in fMRI analyses: pitfalls and recommendations. *Neuroimage* 91, 412–419. doi: 10.1016/j.neuroimage.2013.12.058

Wu, Y. J., and Thierry, G. (2010). Chinese–English bilinguals reading English hear Chinese. *J. Neurosci.* 30, 7646–7651. doi: 10.1523/JNEUROSCI.1602-10.2010

Wu, Y. J., and Thierry, G. (2013). Fast modulation of executive function by language context in bilinguals. *J. Neurosci.* 33, 13533–13537. doi: 10.1523/JNEUROSCI.4760-12.2013

Wu, Y. J., and Thierry, G. (2017). Brain potentials predict language selection before speech onset in bilinguals. *Brain Lang.* 171, 23–30. doi: 10.1016/j.bandl.2017.04.002

Yan, C. G., Wang, X. D., Zuo, X. N., and Zang, Y. F. (2016). DPABI: data processing & analysis for (resting-state) brain imaging. *Neuroinformatics* 14, 339–351. doi: 10.1007/s12021-016-9299-4

Yang, J., Gates, K. M., Molenaar, P., and Li, P. (2015). Neural changes underlying successful second language word learning: an fMRI study. *J. Neurolinguistics* 33, 29–49. doi: 10.1016/j.jneuroling.2014.09.004

Yang, J., and Li, P. (2012). Brain networks of explicit and implicit learning. *PLoS One* 7:e42993. doi: 10.1371/journal.pone.0042993

Zou, L., Abutalebi, J., Zinszer, B., Yan, X., Shu, H., Peng, D., et al. (2012). Second language experience modulates functional brain network for the native language production in bimodal bilinguals. *Neuroimage* 62, 1367–1375. doi: 10.1016/j.neuroimage.2012.05.062

An ERP Investigation of L2–L1 Translation Priming in Adult Learners

Gabriela Meade[1]*, Katherine J. Midgley[2] and Phillip J. Holcomb[2]

[1] Joint Doctoral Program in Language and Communicative Disorders, San Diego State University and University of California, San Diego, San Diego, CA, United States, [2] Department of Psychology, San Diego State University, San Diego, CA, United States

A longstanding debate centers around how beginning adult bilinguals process words in their second language (L2). Do they access the meaning of the L2 words directly or do they first activate the native language (L1) translation equivalents in order to access meaning? To address this question, we used ERPs to investigate how newly learned L2 words influence processing of their L1 translation equivalents. We taught participants the meanings of 80 novel L2 (pseudo)words by presenting them with pictures of familiar objects. After 3 days of learning, participants were tested in a backward translation priming paradigm with a short (140 ms) stimulus onset asynchrony. L1 targets preceded by their L2 translations elicited faster responses and smaller amplitude negativities than the same L1 targets preceded by unrelated L2 words. The bulk of the ERP translation priming effect occurred within the N400 window (350–550 ms), suggesting that the new L2 words were automatically activating their semantic representations. A weaker priming effect in the preceding window (200–350 ms) was found at anterior sites, providing some evidence that the forms of the L1 translation equivalents had also been activated. These results have implications for models of L2 processing at the earliest stages of learning.

Keywords: translation priming, second language acquisition, word learning, lexical mediation, semantic mediation, bilingualism, ERPs

INTRODUCTION

Adult learners of a second language (L2) already have an established system of linguistic and conceptual knowledge in their native language (L1). How L2 words are integrated into that system as they are learned continues to be debated (e.g., Jiang, 2000; Jiang and Forster, 2001; Brysbaert and Duyck, 2010; Grainger et al., 2010; Kroll et al., 2010; Ma et al., 2017; Meade and Dijkstra, 2017). The debate is motivated by leading theories of sequential bilingualism, including the Revised Hierarchical Model (RHM; e.g., Kroll and Stewart, 1994; Kroll et al., 2010) and the Developmental Bilingual Interactive-Activation Model (BIA-d; Grainger et al., 2010), which posit that L2 processing differs as a function of proficiency. At high levels of proficiency, bilinguals are thought to process L1 and L2 words similarly, with direct connections between lexical representations in both languages and a shared semantic store. At earlier stages of proficiency, these models posit that new L2 words are primarily processed via their L1 translation equivalents (i.e., through lexical mediation). However, recent evidence has begun to contradict the latter, suggesting that more direct access to semantics might be established even in low proficiency bilinguals (see, e.g., Duyck and Brysbaert, 2004; Ma et al., 2017; Meade and Dijkstra, 2017). To further investigate whether new L2 words initially activate their L1 translation equivalents or whether they go directly

to semantics, we taught participants a set of 80 L2 words and tested them in a backward (L2–L1) translation ERP priming paradigm with a short stimulus onset asynchrony (SOA).

Evidence for forward and backward translation priming in proficient bilinguals comes from a number of behavioral studies using a primed lexical decision task. L1 target words are classified as real words faster when they are preceded by their L2 translation equivalents than when they are preceded by unrelated L2 words, and the same is true for L2 targets preceded by L1 primes (see, e.g., Duñabeitia et al., 2010b; Wen and van Heuven, 2017, for a meta-analysis). Theoretically, this translation priming could be either due to spreading of activation from the prime word to the form representation of the target word or due to spreading of activation to a shared semantic representation.

Several lines of research converge to suggest that in proficient bilinguals the effect is largely due to facilitated semantic processing. ERPs have played a critical role in pinpointing the semantic nature of translation priming in proficient bilinguals by lending insight into how the priming effect develops over time. Broadly speaking, semantic priming is associated with the N400 and form priming is associated with earlier components, including the P200 and the N250 (e.g., Grainger and Holcomb, 2009; Guo et al., 2012).[1] In a go/no-go lexical decision study with late proficient Russian–English bilinguals, Geyer et al. (2011) time-locked ERPs to L1 and L2 words preceded by the identical word in the same language, by the translation equivalent, or by an unrelated word. In general, words primed by related items elicited smaller negativities than words primed by unrelated items. The identity priming effect in both L1 and L2 began in the earliest window measured (150–300 ms) while the forward and backward translation priming effects were only observed within the N400 window (300–500 ms). The authors interpreted this pattern to suggest that both form and meaning were primed in the identity conditions, whereas only meaning was primed in the translation conditions (see also, Phillips et al., 2006). Duñabeitia et al. (2010a) reported a similar pattern with balanced Basque–Spanish bilinguals in a go/no-go semantic categorization masked priming paradigm; identity priming effects were found within both the N250 and N400 windows, but translation priming effects were restricted to the N400 window. Thus, the timing of the translation priming effect in proficient bilinguals is more consistent with facilitated semantic processing than with facilitated processing of the form of the translation equivalent.

The translation recognition paradigm is another approach to probing the mechanisms that underlie translation priming and, by extension, L2 word processing. In this paradigm, participants see pairs of words and decide whether the two words are correct translations of one another. On critical trials, the L1 target is not the correct translation (e.g., *ajo*) of the L2 prime (e.g., *garlic* for Spanish–English bilinguals), but is related to it either in form (e.g., *ojo* is a form neighbor of *ajo*, but has the semantically unrelated meaning 'eye') or meaning (e.g., cebolla means 'onion';

e.g., Talamas et al., 1999). In proficient bilinguals, both form and semantic distractors produce behavioral interference effects (i.e., slower and less accurate responses compared to unrelated incorrect translations; e.g., Altarriba and Mathis, 1997; Ferré et al., 2006; Moldovan et al., 2016). This suggests that both the meanings and the form of the translation equivalents are activated and make it more difficult to reject the distractors as incorrect translations.

Nevertheless, in proficient bilinguals the behavioral interference effect tends to be larger for semantic distractors than for form distractors, reinforcing that meaning plays a major role in processing of L2 words (e.g., Talamas et al., 1999; Ferré et al., 2006). In an ERP translation recognition task, Guo et al. (2012) also demonstrated that the semantic pathway is more automatic than the lexical pathway in proficient bilinguals. At a 750 ms SOA, form distractors elicited larger amplitude P200s than unrelated targets and semantic distractors elicited smaller amplitude N400s than unrelated targets. Consistent with behavioral results, this pattern suggests that L2 primes were activating both their meanings and the forms of their translation equivalents. However, at a 300 ms SOA, the effect for form distractors within the P200 window disappeared. This prompted the authors to suggest that semantic representations were activated before the L1 translation equivalents (see also, Moldovan et al., 2016). Guo et al.'s (2012) electrophysiological data and SOA manipulations provided detailed time-course information that supports semantics as a primary source of translation priming for proficient bilinguals. This conclusion is consistent with the RHM and the BIA-d in that both models posit direct semantic access for L2 words in proficient bilinguals.

The question that remains unanswered is whether facilitated semantic processing also underlies translation priming in less proficient bilinguals, for whom these theoretical models posit a different lexical architecture. Both the RHM and the BIA-d postulate that L2 words are only directly connected to their L1 translation equivalents in less proficient bilinguals. Therefore, backward translation priming should be lexically mediated, with pre-activation of the form of the L1 translation equivalent as the primary catalyst of the priming effect. Previous empirical studies with less proficient bilinguals have yielded mixed results. For one, it is not clear whether backward translation priming even occurs in the lexical decision task in these less proficient bilinguals (e.g., Jiang and Forster, 2001; Duyck and Warlop, 2009; Dimitropoulou et al., 2011a,b; Witzel and Forster, 2012). For another, the approaches described above to dissociate between the contributions of the semantic representation versus the L1 translation equivalent have been inconclusive.

Early behavioral evidence from translation recognition paradigms was consistent with processing of L2 words via lexical mediation, as proposed in the RHM and BIA-d. The finding of an interference effect for form distractors (e.g., *ojo* instead of *ajo*) was interpreted to suggest that lower proficiency bilinguals were relying on activation of the L1 translation equivalent to process L2 words (e.g., Talamas et al., 1999; Ferré et al., 2006). This argument was especially convincing given the absence of the analogous effect for semantic distractors in the same participants (i.e., no significant differences in response times between semantic

---

[1] Whether P200 or N250 amplitude is reported seems to depend largely on the task that was used. The P200 is often reported in translation recognition studies in which the form-related distractor is predicted to be more difficult to process. In contrast, the N250 is more often reported in translation and identity priming studies in which the related conditions are predicted to be easier to process.

distractors and unrelated targets). In other words, it appeared that the form of the L1 translation equivalent was being activated but the semantic representation was not, perfectly in line with model predictions.

Reports of semantic interference in translation recognition tasks have since challenged the original null semantic finding, indicating that the meanings of the L2 primes can be activated in low proficiency bilinguals under certain conditions (e.g., Sunderman and Kroll, 2006). It is difficult to determine on the basis of these behavioral data alone whether the semantic interference – when it is present – is driven by direct activation of the meaning or indirect semantic activation via the L1 translation equivalent. This is especially true given that these studies used a relatively long 500 ms SOA, which presumably allowed sufficient time for the indirect route. However, recent ERP data from a translation recognition task with lower proficiency bilinguals points to semantics as the primary processing pathway for L2 words (Ma et al., 2017). Similar to proficient bilinguals (Guo et al., 2012), at a 300 ms SOA, there was no effect for form distractors within the P200 window, but there was an effect for semantic distractors within the N400 window. Thus, although translation recognition behavioral data have been inconclusive, recent ERP evidence suggests that L2 comprehension might be semantically mediated, even in unbalanced bilinguals.

Results from standard ERP masked priming paradigms with unbalanced bilinguals also appear more consistent with direct semantic access. For example, Midgley et al. (2009) failed to find evidence of backward translation priming within the N250 window using a 67 ms SOA. The translation priming effect for L1 targets preceded by masked L2 translation primes was restricted to the N400 window, suggesting that the L2 primes were activating their meanings, but not the forms of their L1 translation equivalents. However, in a subsequent masked priming study with slightly more proficient participants and a longer (120 ms) SOA, Schoonbaert et al. (2011) found a widespread N250 priming effect for L1 targets preceded by L2 primes (i.e., smaller negativities for L1 targets in translation pairs compared to those in unrelated pairs). Note that, theoretically, the backward translation N250 priming effect should decrease as proficiency increases and reliance on the L1 translation equivalent diminishes, which is opposite the pattern found across these studies. Instead, the authors suggest that the longer SOA in the study by Schoonbaert et al. (2011) allowed participants to process the L2 primes enough to activate L1 translation equivalents at the form level. Given that backward translation N400 priming effects were robust even at the shorter SOA, these studies seem to suggest that meaning is the primary processing pathway for L2 words, even before high levels of proficiency are achieved.

In interpreting these studies, it is important to keep in mind that the bilingual participants, though unbalanced, had relatively high levels of L2 proficiency. For example, although Midgley et al. (2009) categorized their participants as second language learners, the participants' average self-ratings of L2 language skills were about 4 on a Likert scale from 1 (unable) to 7 (expert). Accurately quantifying the proficiency level of bilinguals who have learned in a classroom and/or immersion setting is

challenging (e.g., Grosjean, 1998) and can differ depending on the measurement tool (e.g., Gollan et al., 2012). At the same time, the proficiency level at which the theoretical transition from lexical mediation to semantic mediation occurs has yet to be specified. Therefore, it remains possible that these participants had already surpassed the proficiency level at which the transition takes place, which would make the evidence of semantic mediation in relatively low proficiency bilinguals more consistent with the theoretical models.

Testing for translation priming effects in the context of a word learning experiment is one way to circumvent this issue of when the transition from lexical mediation to direct semantic access occurs. In fact, deconstructing translation pathways in participants who begin learning their L2 as part of the experiment would seem to be one of the most rigorous tests of the theoretical models that propose lexical mediation at low levels of proficiency. A handful of such priming studies with learners have been conducted, but have failed to yield conclusive results thus far (e.g., Altarriba and Mathis, 1997; Mestres-Missé et al., 2007; Dobel et al., 2009; Witzel and Forster, 2012; Pu et al., 2016). For example, after teaching English monolinguals a set of 36 Spanish words, Altarriba and Mathis (1997) found behavioral interference effects for both form and semantic distractors in a translation recognition task with a 300 ms SOA. This suggests that both the L1 translation equivalent forms and the meanings of the new L2 words were activated. Emerging ERP evidence supports the claim that L2 words activate both their meanings and the forms of their translation equivalents in learners (e.g., Mestres-Missé et al., 2007; Pu et al., 2016). For example, Pu et al. (2016) taught native English speakers 112 Spanish words through explicit paired associations (e.g., *cama-bed*) and tested them in translation verification task (i.e., are these word pairs correct translations?). They found that targets in translation pairs elicited smaller negativities than targets in unrelated pairs (i.e., priming) beginning between 200 and 300 ms and continuing through the N400 window. The authors interpret the early onset of the priming effect as support for lexically mediated backward translation, but acknowledge that the 800 ms SOA was long enough to allow for strategic activation of the form of the L1 translation. Especially in light of recent findings that the duration of the SOA influences ERP priming patterns in low proficiency bilinguals (Ma et al., 2017), it is important to test whether or not these priming patterns hold at a short SOA.

In summary, there is robust evidence for N400 effects in translation priming studies, which supports semantic mediation among bilinguals at all levels of proficiency (e.g., Midgley et al., 2009; Duñabeitia et al., 2010a; Geyer et al., 2011; Schoonbaert et al., 2011; Guo et al., 2012; Pu et al., 2016; Ma et al., 2017). The evidence for earlier, form-based priming effects is comparatively limited and is mostly observed in studies with long SOAs (e.g., Guo et al., 2012; Pu et al., 2016; Ma et al., 2017). However, almost all of these studies have been done with bilinguals who had already achieved some L2 proficiency. To further investigate whether L2 words are processed via lexical mediation at the earliest stages of learning, we taught participants a set of novel L2 words and tested them in a backward priming paradigm with a 140 ms SOA. The L2 words were initially pseudowords that

were paired with pictures representing their meanings during the learning phase of the experiment. Following Pu et al. (2016), after learning we recorded EEG as participants saw L2 prime – L1 target pairs and decided whether the two words were correct translations or not. Using a shorter SOA than in the study by Pu et al. (2016) allowed us to minimize overt translation and index the representations that are automatically and rapidly activated during processing of newly learned L2 words. We predicted that L1 targets preceded by their L2 translations would elicit faster responses and smaller amplitude negativities (i.e., priming) than L1 targets preceded by unrelated L2 words. As argued above, the onset of ERP effects is critical for determining whether translation priming is lexically or semantically motivated in these learners. A priming effect solely within the N400 window would be consistent with activation of the semantic representation. Finding an effect before the N400, in time windows that are commonly associated with processing of lexical form (i.e., P200/N250), would suggest that the form of the L1 translation had been activated during processing of the L2 prime. The latter would be consistent with the lexical mediation posited for low proficiency bilinguals in the RHM and BIA-d.

MATERIALS AND METHODS

Participants

Participants included 18 young adults who were right-handed and had normal or corrected-to-normal vision. By self-report, they were not fluent in any language other than English and were not exposed to any language other than English before the age of 6. Participants reported having no history of neurological dysfunction or language disorders, and were not taking any medications that would affect brain function. Data from these same participants in other tasks have been reported elsewhere (Meade et al., 2018). In addition to the three participants excluded from the original report, data from two additional participants were excluded here. One participant was excluded for high artifact rejection due to blinks in this task (>30% of trials) and the other was the participant with the lowest overall accuracy in the priming task that could be rejected to maintain the counterbalancing described below.[2]

Stimuli

Stimuli included 86 L2 words (80 critical items and 6 practice items) that were drawn from the ARC Nonword Database (Rastle et al., 2002) and chosen to be orthotactically and phonotactically legal in English (e.g., *grif*, *labe*, *slont*, and *plurd*). All of the L2 words were four to five letters in length; more characteristics of these L2 words and their L1 translations can be found in **Table 1**. During the learning exercises, the L2 words were paired with pictures depicting familiar objects. All of the pictures had naming agreement at or above 85% in previous norming studies (mean = 97%; Bates et al., 2003). Form overlap between the L2 words and their L1 translation equivalents was minimized.

The average Levenshtein distance (i.e., number of insertions, deletions, and substitutions) between the L2 words and their L1 translations was 5.12 (SD: 1.20). A full description of the L2 words can be found in Meade et al. (2018).

Procedure

Participants were instructed that they would be learning words from a new language. In order to reinforce that these were words from another language, the experiment began with a language decision ERP pretest (i.e., press one button for English words and another button for words from another language; see Meade et al., 2018). Learning exercises were then administered over three consecutive days beginning the day of the pretest. Each word was presented a total of 12 times during the learning phase, either in the context of a two-alternative forced-choice (2AFC) task or a typing task (see **Table 2**). In the 2AFC task, a picture was presented with two L2 words and participants had to choose which of the L2 words matched the picture. Participants received feedback after each trial in which the picture was displayed together with the correct L2 word. In the typing task, they saw the picture and had to type the corresponding L2 word. If they typed the correct word, they moved on to the subsequent trial. If they typed the incorrect word, the correct word was displayed and they were then asked to type the correct word. On Day 1 of training, they had the first and last letters of the word as a cue in the typing task, but by the last session they had no cues (see **Table 2**). By the last learning session, mean accuracy was 99% (SD: 1.2%) in the 2AFC task and 95% (SD: 4.3%) in the typing task, which demonstrates that the participants had successfully learned the words.

On the fourth day of the experiment, participants took part in an ERP post-test that included a backward priming paradigm. An L2 prime was presented in lowercase for 140 ms, followed

TABLE 1 | Stimulus characteristics [mean (SD)].

	Length	N	Frequency	Concreteness
L2 words	4.50	4.35	–	–
	(0.50)	(4.42)		
L1 translations	5.39	4.98	79.69	4.54
	(1.64)	(5.63)	(252.81)	(0.33)

N = the number of orthographic neighbors in English. N and frequency were extracted from the MCWord database (Medler and Binder, 2005). Concreteness ratings (on a scale from 1 = abstract to 5 = concrete) are from Brysbaert et al. (2014).

TABLE 2 | Training overview.

Day 1	Day 2	Day 3
Paired associate/2AFC	2AFC	2AFC
2AFC	Typing[‡]	Typing[†]
Typing[‡]	2AFC	2AFC
	Typing[†]	Typing

2AFC = Two-alternative forced-choice. [‡]First and last letters provided as a cue. [†]First letter provided as a cue. Data from these tasks are available in Meade et al. (2018).

---

[2]Note that analyzing the data from all 20 participants who were included the original report yields the same pattern of results that we report here.

immediately by an L1 target in uppercase that remained on the screen for 500 ms. Participants were asked to decide as quickly as possible whether the two words were correct translations and to press one button if they were and another button if they were not. Response hand was counterbalanced across participants. One thousand ms after the response, a purple fixation cross appeared, during which participants were instructed to blink if needed. After 1500 ms, the purple fixation cross turned white for 900 ms and then a 500 ms blank screen signaled the beginning of the next trial. Before beginning the experiment, there was a practice that included three translation and three unrelated trials, none of which were included in the actual experiment.

Each L2 word was presented twice, followed by the correct L1 translation in one half of the experiment and by an unrelated L1 word in the other half. All participants saw the same list (e.g., *grif-ORANGE* in the first half and *grif-KNIGHT* in the second half). However, the pairings between the words and pictures during the learning phase were systematically controlled across participants such that any given pair was the correct translation for half of the participants and unrelated for the other half (e.g., nine participants learned the L2 word *grif* with a picture of an orange and nine of them learned the L2 word *grif* with a picture of a knight). This design ensured that the same L1 targets occurred in the translation and unrelated conditions.
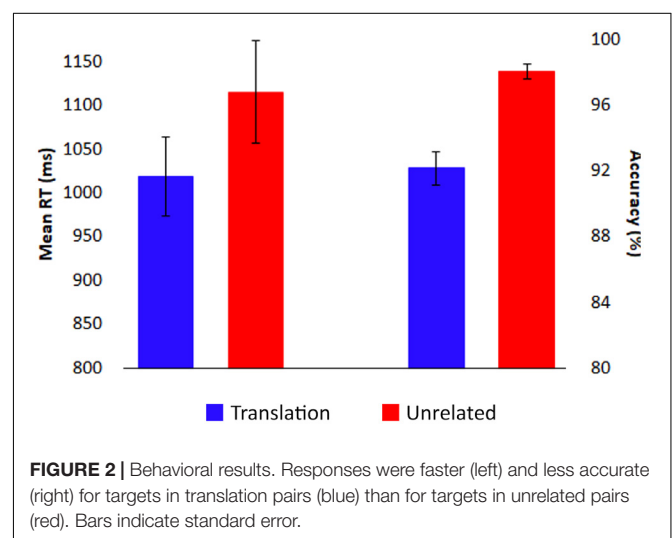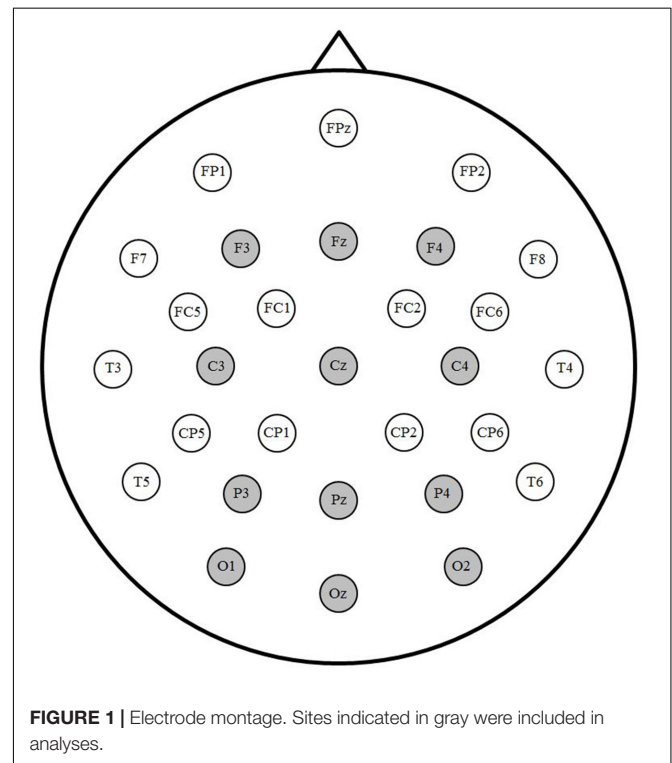
EEG Recording and Analysis

EEG was recorded from 29 electrodes in an Electro-Cap using a left mastoid reference. It was amplified with SynAmpsRT amplifiers (Neuroscan-Compumedics) using a band pass of DC to 100 Hz and was sampled continuously at 500 Hz. Off-line, ERPs were time-locked to target onset for each participant and prime condition (translation and unrelated) separately using a 100 ms pre-stimulus baseline and a 15 Hz low-pass filter. A loose electrode placed below the left eye was used in conjunction with recordings from FP1 to detect blink artifacts and another electrode on the outer canthus of the right eye was used to detect horizontal eye movements. Impedances were maintained below 10 kΩ for eye electrodes and below 2.5 kΩ for scalp and reference electrodes. Trials with artifacts during the baseline period or within 1000 ms of target onset were excluded from analyses, as were trials with incorrect responses. In the final analyses, an average of 72 and 76 trials (out of 80) were included in the translation and unrelated conditions, respectively.

A subset of 12 electrodes was selected for statistical analyses (see **Figure 1**). To test for a translation priming effect, ANOVAs with factors Prime (translation and unrelated), Laterality (left, midline, and right) and Anterior/Posterior (frontal, central, parietal, and occipital) were used on mean amplitude within two successive windows. N400 amplitude was measured between 350 and 550 ms, consistent with previous priming studies (e.g., Grainger et al., 2006; Phillips et al., 2006). Due to the short SOA, processing of the prime and target overlapped in time and the morphology of the waveform differed from standard ERPs to single words. The early window (200–350 ms) was chosen based on visual inspection of the grand averaged waveforms to encompass the negative peak preceding the N400.

RESULTS

Response times shorter than 200 ms and longer than 2000 ms were excluded from analyses. As predicted, correct responses were faster for translation trials (mean: 1019 ms) than unrelated trials (mean: 1115 ms), $F(1,17) = 22.70$, $p < 0.001$, $\eta_p^2 = 0.57$ (see **Figure 2**). However, accuracy was slightly higher for unrelated trials (mean: 98%) than for translation trials (mean: 92%), $F(1,17) = 38.86$, $p < 0.001$, $\eta_p^2 = 0.70$, potentially indicative of a speed-accuracy trade-off.

FIGURE 1 | Electrode montage. Sites indicated in gray were included in analyses.

FIGURE 2 | Behavioral results. Responses were faster (left) and less accurate (right) for targets in translation pairs (blue) than for targets in unrelated pairs (red). Bars indicate standard error.

In the ERP analyses, the main effect of Prime was not significant between 200 and 350 ms, $F(1,17) = 0.88$, $p = 0.361$, $\eta_p^2 = 0.05$. However, an interaction between Prime and Anterior/Posterior indicated that the effect went in the expected direction across anterior sites and in the opposite direction across the most posterior electrodes, Prime × Anterior/Posterior, $F(3,51) = 5.73$, $p = 0.021$, $\eta_p^2 = 0.25$ (see **Figure 3**). Follow-up analyses including only the most anterior electrodes (F3, Fz, and F4) confirmed that the priming effect was reliable at those sites, Prime, $F(1,17) = 5.09$, $p = 0.038$, $\eta_p^2 = 0.23$. A point-by-point time course analysis (see **Figure 4**) was consistent in suggesting that there was a weak early effect across frontal sites, but that the most reliable effect began within the N400 window, around 400 ms. Indeed, there was a widespread effect of priming within the N400 window (350–550 ms) that was strongest at central midline sites, Prime, $F(1,17) = 26.88$, $p < 0.001$, $\eta_p^2 = 0.61$, Prime × Laterality, $F(2,34) = 4.34$, $p = 0.029$, $\eta_p^2 = 0.20$, Prime × Anterior/Posterior, $F(3,51) = 4.13$, $p = 0.043$, $\eta_p^2 = 0.19$, Prime × Laterality × Anterior/Posterior, $F(6,102) = 4.51$, $p = 0.004$, $\eta_p^2 = 0.21$ (refer to **Figures 3**, **4**).

DISCUSSION

Leading models of sequential bilingualism, including the RHM and BIA-d, posit that L2 words are processed via their L1 translation equivalents (i.e., lexical mediation) at low levels of proficiency. In contrast, there is growing empirical evidence to suggest that L2 words might be processed directly for meaning at relatively early stages of proficiency. To address this debate, we taught participants novel L2 words and tested them in a backward (L2–L1) priming paradigm with a short (140 ms) SOA. L1 targets in translation pairs elicited faster responses than the same targets in unrelated pairs, indicating that participants had learned the words and were processing them efficiently. ERP effects began as early as 200 ms after target onset at anterior sites, in a window that roughly corresponds to the N250. Such an early effect would appear to be consistent with lexical mediation and pre-activation of the lexical form of the L1 translation equivalent in these learners. However, the bulk of the observed ERP priming effects occurred within the N400 window, which suggests that L2 words were also directly activating their meanings. Given the short SOA, these results suggest that both L1 form and meaning representations were automatically accessed, but to different degrees. With a focus on the relative strength of the priming effects in the two windows, we discuss several potential lexical architectures that could underlie these results.

The early priming effects that we observed differ from typical N250 priming effects, which begin earlier and have a broader distribution. However, there are previous reports of an anterior N250 effect that more closely resembles the one we observed here. In particular, Grainger et al. (2006) found that orthographic overlap between visually-presented primes and targets modulates a posterior N250 whereas phonological overlap modulates a more anterior N250. In light of those results, one potential explanation for the anterior distribution of the early priming effect here is that (only) the phonological forms of the L1 words were primed. This

makes sense given that participants learned the L2 words with pictures that they could name in their L1, but they never saw the orthographic forms of the L1 translation equivalents (until the ERP translation priming paradigm). If this interpretation is correct, it follows that when the orthographic forms of the L1 translations are presented during learning, the distribution of the early priming effects should include a more posterior component. Indeed, after teaching L2 words through lexical association, Pu et al. (2016) found early translation priming effects that appear to have a broader distribution than the effects that we observed here. Directly comparing the early priming effects for L2 words learned with L1 translations versus pictures in future studies would confirm that learning method influences the nature of the L1 form representations that are activated by L2 words.

In contrast, in a translation recognition paradigm with a 300 ms SOA, Ma et al. (2017) did not find a significant effect of form distractors within their P200 window (150–300 ms) and concluded that L1 translations are not automatically activated in low proficiency bilinguals. Several differences between the two studies could explain these divergent results. For one, our participants mastered a small set of L2 words in the context of this study whereas the participants in the study by Ma et al. (2017) were classroom learners and were therefore exposed to a wider range of L2 words in a variety of learning situations. How they learned the words may have affected the strength of activation of L1 translation equivalents. There were also methodological differences between the two studies that could help explain the results. For example, the translation recognition task that Ma et al. (2017) used only indirectly indexes activation of the translation equivalent; responses are recorded to neighbors of the L1 translations (i.e., form distractors) rather than to the L1 translations themselves. Therefore, it is possible that the L1 translation equivalents were also activated in that study, but not enough to interfere with processing of the neighbors. This seems especially plausible since the form priming effects that we observed were on the smaller side. The different SOAs between the two studies also likely influenced the results. It could be that activation of the L1 translation equivalent is transient such that it was strong enough to be measured at the 140 ms SOA here, but did not persist through the 300 ms SOA in the study by Ma et al. (2017). Evidence from priming studies with monolinguals supports this hypothesis; N250 (but not N400) effects become refractory at SOAs of 300 ms or more (Holcomb and Grainger, 2007). It is also important to note with a 140 ms SOA, the N250 window that we measured (200–350 ms after target onset) is temporally congruent with the N400 elicited by the primes (340–490 ms after prime onset). Some portion of this effect could therefore be driven by backward semantic priming from the L1 target to the L2 prime. More research is needed to test the effect of SOA in translation priming studies and, more generally, to determine which of these design differences led to the early priming effect here.

Although we found evidence of lexical mediation, the relative difference in size of the early lexical effect and the later N400 effect suggests that direct semantic activation was likely producing much of the priming effect. The typical centro-posterior distribution of our N400 effect suggests that
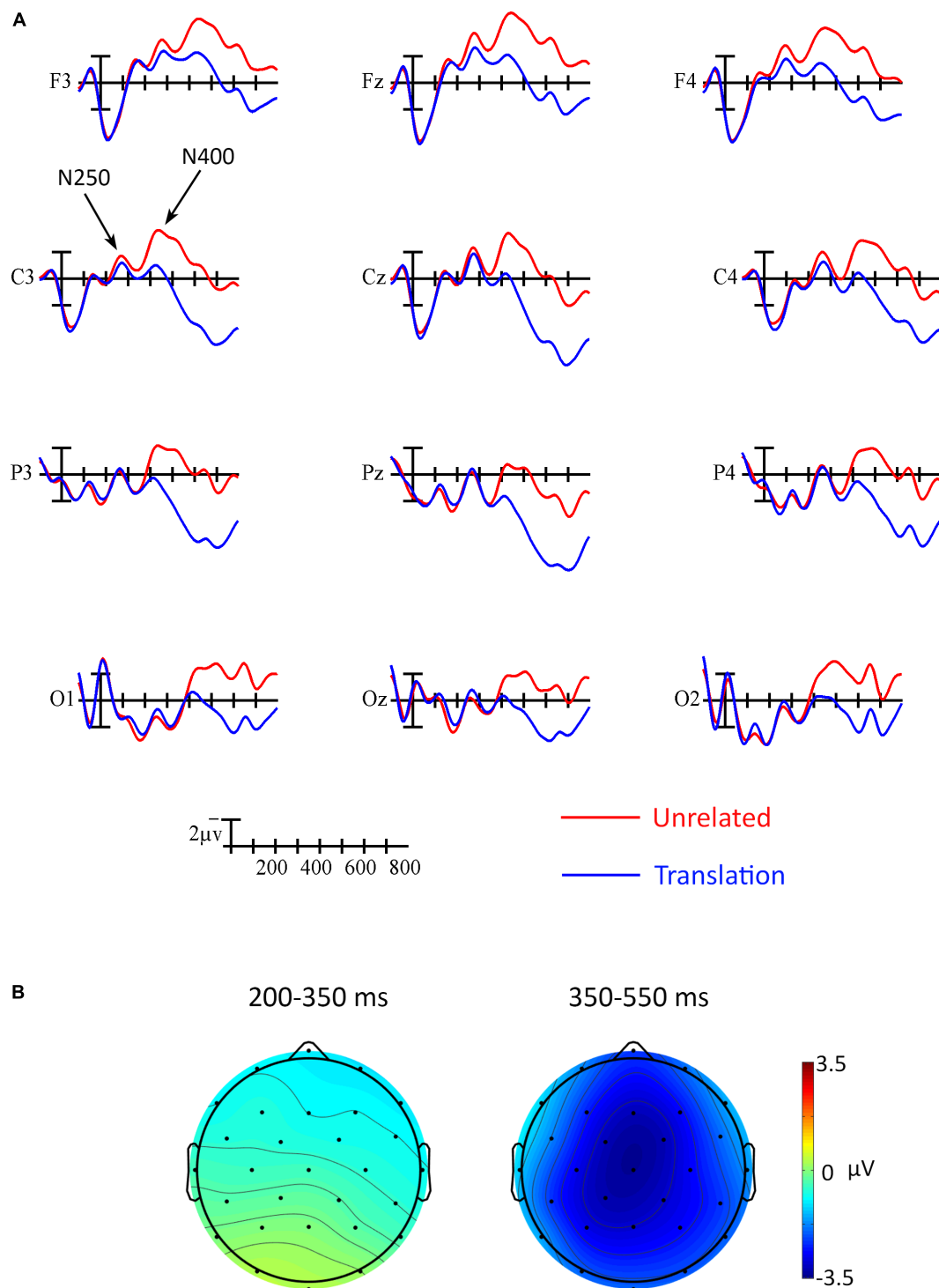
FIGURE 3 | ERP results. **(A)** Grand averaged ERP waveforms elicited by targets in translation pairs (blue) and unrelated pairs (red). Each vertical tick marks 100 ms, the calibration bar marks 2 μV, and negative is plotted up. The 250 and N400 are indicated at site C3. **(B)** Scalp voltage maps showing the effect of translation priming (unrelated-translation) for each of the analyzed time windows.

it resulted from spreading of activation within the semantic system. This contrasts with the fronto-central N400 priming effect that Mestres-Missé et al. (2007) found in a 500 ms

SOA backward priming paradigm with learners. In that study, participants implicitly learned the meanings of novel words presented at the end of three L1 sentences with increasing
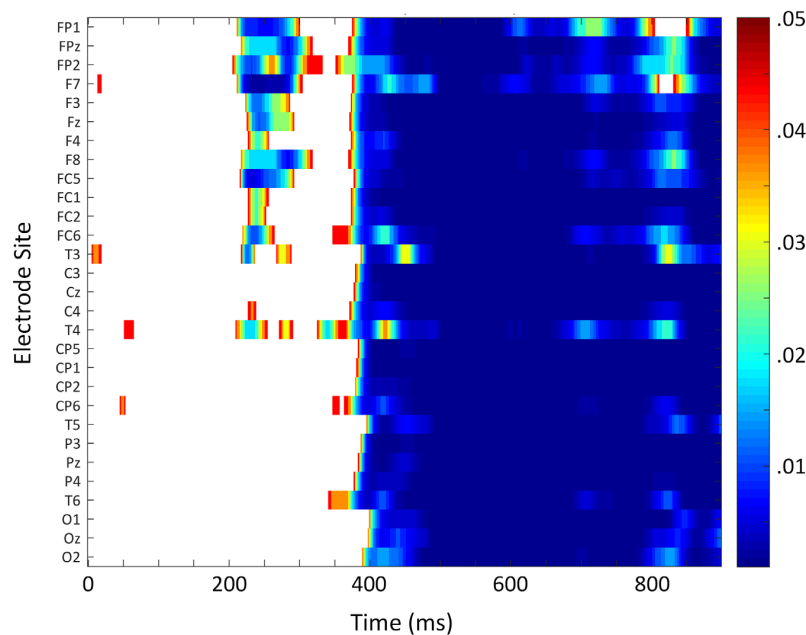
FIGURE 4 | Time course analysis of ERP translation priming effect. False discovery rate-corrected *p*-values for each time point at each electrode site. Color indicates where the priming translation effect was significant.

contextual constraint and were tested in the priming paradigm the same day. The authors attributed the frontal distribution of the priming effect to recruitment of prefrontal regions and an increase in cognitive control during semantic retrieval of the new words. These two studies might represent two different stages of L2 learning as described, for example, in the episodic hypothesis of L2 learning (e.g., Jiang, 2000; Jiang and Forster, 2001; Witzel and Forster, 2012). Proponents of the episodic hypothesis differentiate between "lexical knowledge," which involves storing information about L2 words in general episodic memory and "lexical competence," which denotes that lexicosemantic information has been integrated into the linguistic system. The frontal N400 effect reported by Mestres-Missé et al. (2007) could be indicative of the controlled meaning retrieval that characterizes the lexical knowledge stage, whereas the more typical N400 distribution that we observed in the present study suggests that L2 words can be integrated into the lexicosemantic system over a span of only a few days.

How do we account for activation of both the L1 translation equivalent and the meaning? It would appear that these data reflect a combination of lexical mediation and semantic mediation, or the transition from one to the other. In both the RHM and the BIA-d, the lexical links decrease in strength as proficiency increases and direct semantic links are established, but they never disappear entirely. Thus, it is possible that these words were being primarily processed through the semantic route, but residual activation was also flowing to the L1 translation equivalent via the weakened lexical links. It could also be that individual L2 words were at different stages of the transition from lexical mediation to semantic mediation

(see, e.g., Kroll and Tokowicz, 2005). In other words, the two patterns in the averaged ERPs might reflect processing via lexical mediation for a (small) subset of the L2 words and processing via semantic mediation for a (larger) subset of the L2 words. In the BIA-d, this transition is implemented by decreasing the lexical "clamping" between each L2 word and its translation equivalent and increasing top–down inhibition of the L1 from the L2 language node (Grainger et al., 2010). We know from studies with proficient bilinguals that the translation priming effect should onset within the N400 window in the final state (e.g., Phillips et al., 2006; Duñabeitia et al., 2010b; Geyer et al., 2011). This could be achieved either by further weakening of the lexical links for all words or by processing a larger majority of the L2 words via semantic mediation. If the latter is true, and the transition is happening at the level of individual words, it would be informative to know what lexicosemantic characteristics allow certain words to transition faster than others.

CONCLUSION

The present study offers new evidence for both early (N250-like) and later (N400) translation priming effects at a short SOA that precludes strategic processing. The N400 priming effect was substantially larger than the earlier anterior effect. It is therefore unlikely that it resulted purely from the indirect (i.e., lexically mediated) semantic processing posited in the RHM and BIA-d. Rather, the data are more consistent with direct semantic access after relatively few exposures to new L2 words. Whether all of the L2 words were being processed via this direct semantic pathway is not clear. The early form priming effects could be due to weak

activation of the L1 translation equivalents of all L2 words or, alternatively, to strong activation of the L1 translation equivalents of a small subset of L2 words that were still being processed via lexical mediation. How these dynamics would differ among classroom students who learn a more diverse set of words as part of a more ecologically valid language learning experience also remains unknown. Tracking the relative contributions of lexical versus semantic mediation over the course of learning, including in adults who learn their L2 in more typical classroom settings, will begin to answer these important questions.

ETHICS STATEMENT

All participants gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Institutional Review Board at San Diego State University.

AUTHOR CONTRIBUTIONS

GM and PH designed the study. GM collected and analyzed the data. GM, KM, and PH participated in writing the paper.

FUNDING

REFERENCES

Altarriba, J., and Mathis, K. M. (1997). Conceptual and lexical development in second language acquisition. *J. Mem. Lang.* 36, 550–568. doi: 10.1006/jmla.1997.2493

Bates, E., D'Amico, S., Jacobsen, T., Székely, A., Aandonova, E., Devescovi, A., et al. (2003). Timed picture naming in seven languages. *Psychon. Bull. Rev.* 10, 344–380. doi: 10.3758/BF03196494

Brysbaert, M., and Duyck, W. (2010). Is it time to leave behind the Revised Hierarchical Model of bilingual language processing after fifteen years of service? *Bilingualism* 13, 359–371. doi: 10.1017/S1366728909990344

Brysbaert, M., Warriner, A. B., and Kuperman, V. (2014). Concreteness rating for 40 thousand generally known English word lemmas. *Behav. Res. Methods* 46, 904–911. doi: 10.3758/s13428-013-0403-5

Dimitropoulou, M., Duñabeitia, J. A., and Carreiras, M. (2011a). Masked translation priming effects with low proficient bilinguals. *Mem. Cogn.* 39, 260–275. doi: 10.3758/s13421-010-0004-9

Dimitropoulou, M., Duñabeitia, J. A., and Carreiras, M. (2011b). Two words, one meaning: evidence of automatic co-activation of translation equivalents. *Front. Psychol.* 2:188. doi: 10.3389/fpsyg.2011.00188

Dobel, C., Junghöfer, M., Breitenstein, C., Klauke, B., Knecht, S., Pantev, C., et al. (2009). New names for known things: on the association of novel word forms with existing semantic information. *J. Cogn. Neurosci.* 22, 1251–1261. doi: 10.1162/jocn.2009.21297

Duñabeitia, J. A., Dimitropoulou, M., Uribe-Etxebarria, O., Laka, I., and Carreiras, M. (2010a). Electrophysiological correlates of the masked translation priming effect with highly proficient simultaneous bilinguals. *Brain Res.* 1359, 142–154. doi: 10.1016/j.brainres.2010.08.066

Duñabeitia, J. A., Perea, M., and Carreiras, M. (2010b). Masked translation priming effects with highly proficient bilinguals. *Exp. Psychol.* 57, 98–107. doi: 10.1027/1618-3169/a000013

Duyck, W., and Brysbaert, M. (2004). Forward and backward number translation requires conceptual mediation in both balanced and unbalanced bilinguals. *J. Exp. Psychol. Hum. Percept. Perform.* 30, 889–906. doi: 10.1037/0096-1523.30.5.889

Duyck, W., and Warlop, N. (2009). Translation priming between the native language and a second language: new evidence from Dutch-French bilinguals. *Exp. Psychol.* 56, 173–179. doi: 10.1027/1618-3169.56.3.173

Ferré, P., Sánchez-Casas, R., and Guasch, M. (2006). Can a horse be a donkey? Semantic and form interference effects in translation recognition in early and late proficient and nonproficient Spanish-Catalan bilinguals. *Lang. Learn.* 56, 571–608. doi: 10.1111/j.1467-9922.2006.00389.x

Geyer, A., Holcomb, P. J., Midgley, K. J., and Grainger, J. (2011). Processing words in two languages: an event-related potential study of proficient bilinguals. *J. Neurolinguist.* 24, 338–351. doi: 10.1016/j.jneuroling.2010.10.005

Gollan, T. H., Weissberger, G. H., Runnqvist, E., Montoya, R. I., and Cera, C. M. (2012). Self-ratings of spoken language dominance: a multi-lingual naming test (MINT) and preliminary norms for young and aging Spanish-English bilinguals. *Bilingualism* 15, 594–615. doi: 10.1017/S1366728911000332

Grainger, J., and Holcomb, P. J. (2009). Watching the word go by: on the time-course of component processes in visual word recognition. *Lang. Linguist. Compass* 3, 128–156. doi: 10.1111/j.1749-818X.2008.00121.x

Grainger, J., Kiyonaga, K., and Holcomb, P. J. (2006). The time course of orthographic and phonological code activation. *Psychol. Sci.* 17, 1021–1026. doi: 10.1111/j.1467-9280.2006.01821.x

Grainger, J., Midgley, K., and Holcomb, P. J. (2010). "Re-thinking the bilingual interactive-activation model from a developmental perspective," in *Language Acquisition Across Linguistic and Cognitive Systems*, eds M. Kail and M. Hickmann (Amsterdam: John Benjamins Publishing Company), 267–283. doi: 10.1075/lald.52.18gra

Grosjean, F. (1998). Studying bilinguals: methodological and conceptual issues. *Bilingualism* 1, 131–149. doi: 10.1017/S136672899800025X

Guo, T., Misra, M., Tam, J. W., and Kroll, J. F. (2012). On the time course of accessing meaning in a second language: an electrophysiological and behavioral investigation of translation recognition. *J. Exp. Psychol. Lear. Mem. Cogn.* 38, 1165–1186. doi: 10.1037/a0028076

Holcomb, P. J., and Grainger, J. (2007). Exploring the temporal dynamics of visual word recognition in the masked repetition priming paradigm using event-related potentials. *Brain Res.* 1180, 39–58. doi: 10.1016/j.brainres.2007.06.110

Jiang, N. (2000). Lexical representation and development in a second language. *Appl. Linguist.* 21, 47–77. doi: 10.1093/applin/21.1.47

Jiang, N., and Forster, K. I. (2001). Cross-language priming asymmetries in lexical decision and episodic recognition. *J. Mem. Lang.* 44, 32–51. doi: 10.1006/jmla.2000.2737

Kroll, J. F., and Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connection between bilingual memory representations. *J. Mem. Lang.* 33, 149–174. doi: 10.1006/jmla.1994.1008

Kroll, J. F., and Tokowicz, N. (2005). "Models of bilingual representation and processing: looking back and to the future," in *Handbook of Bilingualism: Psycholinguistic Approaches*, eds J. F. Kroll and A. M. B. de Groot (New York, NY: Oxford University Press), 531–553.

Kroll, J. F., Van Hell, J. G., Tokowicz, N., and Green, D. W. (2010). The revised hierarchical model: a critical review and assessment. *Bilingualism* 13, 373–381. doi: 10.1017/s136672891000009x

Ma, F., Chen, P., Guo, T., and Kroll, J. F. (2017). When late second language learners access the meaning of L2 words: using ERPs to investigate the role of the L1 translation equivalent. *J. Neurolinguist.* 41, 50–69. doi: 10.1016/j.jneuroling.2016.09.006

Meade, G., and Dijkstra, T. (2017). "Mechanisms underlying word learning in second language acquisition," in *Bilingualism: A Framework for Understanding*

the Mental Lexicon, eds M. Libben, G. Libben, and M. Goral (Philadelphia, PA: John Benjamins Publishing Company), 49–72. doi: 10.1075/bpa.6.03mea

Meade, G., Midgley, K. J., Dijkstra, T., and Holcomb, P. J. (2018). Cross-language neighborhood effects in learners indicative of an integrated lexicon. *J. Cogn. Neurosci.* 30, 70–85. doi: 10.1162/jocn_a_01184

Medler, D. A., and Binder, J. R. (2005). *MCWord: An On-Line Orthographic Database of the English Language*. Available at: http://www.neuro.mcw.edu/mcword

Mestres-Missé, A., Rodriguez-Fornells, A., and Münte, T. F. (2007). Watching the brain during meaning acquisition. *Cereb. Cortex* 17, 1858–1866. doi: 10.1093/cercor/bhl094

Midgley, K. J., Holcomb, P. J., and Grainger, J. (2009). Masked repetition and translation priming in second language learners: a window on the time course of form and meaning activation using ERPs. *Psychophysiology* 46, 551–565. doi: 10.1111/j.1469-8986.2009.00784.x

Moldovan, C. D., Demestre, J., Ferré, P., and Sánchez-Casas, R. (2016). The role of meaning and form similarity in translation recognition in highly proficient balanced bilinguals: a behavioral and ERP study. *J. Neurolinguist.* 37, 1–11. doi: 10.1016/j.jneuroling.2015.07.002

Phillips, N. A., Klein, D., Mercier, J., and de Boysoon, C. (2006). ERP measures of auditory word repetition and translation priming in bilinguals. *Brain Res.* 1125, 116–131. doi: 10.1016/j.brainres.2006.10.002

Pu, H., Holcomb, P. J., and Midgley, K. (2016). Neural changes underlying early stages of L2 vocabulary acquisition. *J. Neurolinguist.* 40, 55–65. doi: 10.1016/j.jneuroling.2016.05.002

Rastle, K., Harrington, J., and Coltheart, M. (2002). 358,534 nonwords: the ARC nonword database. *Q. J. Exp. Psychol. A Hum. Exp. Psychol.* 55, 1339–1362. doi: 10.1080/02724980244000099

Schoonbaert, S., Holcomb, P. J., Grainger, J., and Hartsuiker, R. J. (2011). Testing asymmetries in noncognate translation priming: evidence from RTs and ERPs. *Psychophysiology* 48, 74–81. doi: 10.1111/j.1469-8986.2010.01048.x

Sunderman, G., and Kroll, J. F. (2006). First language activation during second language lexical processing: an investigation of lexical form, meaning, and grammatical class. *Stud. Second Lang. Acquis.* 28, 387–422. doi: 10.1017/S0272263106060177

Talamas, A., Kroll, J. F., and Dufour, R. (1999). From form to meaning: stages in the acquisition of second-language vocabulary. *Bilingualism* 2, 45–58. doi: 10.1017/S1366728999000140

Wen, Y., and van Heuven, W. J. B. (2017). Non-cognate translation priming in masked priming lexical decision experiments: a meta-analysis. *Psychon. Bull. Rev.* 24, 879–886. doi: 10.3758/s13423-016-1151-1

Witzel, N. O., and Forster, K. I. (2012). How L2 words are stored: the episodic L2 hypothesis. *J. Exp. Psychol. Learn. Mem. Cogn.* 38, 1608–1621. doi: 10.1037/a0028072

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.

Check for
updates

The Influence of Concreteness of Concepts on the Integration of Novel Words into the Semantic Network

Jinfeng Ding [1,2], Wenjuan Liu [1,2] and Yufang Yang [1,2]*

[1] *CAS Key Laboratory of Behavioral Science, Institute of Psychology, Beijing, China,* [2] *Department of Psychology, University of Chinese Academy of Sciences, Beijing, China*

On the basis of previous studies revealing a processing advantage of concrete words over abstract words, the current study aimed to further explore the influence of concreteness on the integration of novel words into semantic memory with the event related potential (ERP) technique. In the experiment during the learning phase participants read two-sentence contexts and inferred the meaning of novel words. The novel words were two-character non-words in Chinese language. Their meaning was either a concrete or abstract known concept which could be inferred from the contexts. During the testing phase participants performed a lexical decision task in which the learned novel words served as primes for either their corresponding concepts, semantically related or unrelated targets. For the concrete novel words, the semantically related words belonged to the same semantic categories with their corresponding concepts. For the abstract novel words, the semantically related words were synonyms of their corresponding concepts. The unrelated targets were real words which were concrete or abstract for the concrete or abstract novel words respectively. The ERP results showed that the corresponding concepts and the semantically related words elicited smaller N400s than the unrelated words. The N400 effect was not modulated by the concreteness of the concepts. In addition, the concrete corresponding concepts elicited a smaller late positive component (LPC) than the concrete unrelated words. This LPC effect was absent for the abstract words. The results indicate that although both concrete and abstract novel words can be acquired and linked to their related words in the semantic network after a short learning phase, the concrete novel words are learned better. Our findings support the (extended) dual coding theory and broaden our understanding of adult word learning and changes in concept organization.

Keywords: concreteness, novel word learning, context, semantic memory, ERP

INTRODUCTION

A processing advantage of concrete concepts over abstract concepts, namely a concreteness effect, is reported in a variety of tasks including lexical decision, free recall, recognition, as well as paired associate learning (for reviews, see Paivio, 1991; Schwanenflugel et al., 1992). Concreteness effects are mainly explained by dual coding theory (Paivio, 1986) or context availability hypothesis (Schwanenflugel and Shoben, 1983).

According to the dual coding theory, there is a verbal-based system and an imagery-based system associated with concepts in semantic memory. The former is responsible for the representation and processing of linguistic information, and the latter for nonverbal information. Concrete words are connected to both the systems, while abstract words are only connected to the verbal system. When a concrete word is processed, the verbal and nonverbal systems function independently and are interconnected, which results in an additive effect, thereby yielding a processing advantage for concrete words over abstract words.

Different from the dual coding theory which emphasize that the representations of concrete and abstract words differ qualitatively, the context availability hypothesis proposes that they differ quantitively. It posits that concreteness effects arise from the differences in availability of contextual information. The contextual information can be retrieved from the person's prior knowledge or from the circumstance in which the stimulus appears. Since people encounter abstract words in a wide range of contexts or circumstances, the contextual information for abstract words is represented in a looser way. Therefore, the poor performance for abstract words is not because of lesser availability of imagery, but because of the relative unavailability of associated contextual information in memory for abstract words (Schwanenflugel et al., 1992).

Later, Holcomb et al. (1999) extended the dual coding theory with the context availability theory. They proposed that the concreteness effect could be attributed to both superior associate connections in the verbal-based system and the use of imagery-based system for concrete words. This extended dual coding theory is supported by many studies (e.g., Jessen et al., 2000; West and Holcomb, 2000; Binder et al., 2005; Zhang et al., 2006). For instance, Jessen et al. (2000) found greater activation in the lower right and left parietal lobes, as well as in the left inferior frontal lobe and in the precuneus for concrete words relative to abstract words. The stronger activation in the left parietal and frontal areas indicated greater verbal context resources for concrete words, and in the right parietal lobe indicated an additional involvement of spatial imagery-based system for concrete words.

The theories mentioned above account for the different representations between concrete and abstract words. These distinctions might reflect the way in which the words have been learned (Mestres-Missé et al., 2014). As expected, the concreteness of concepts not only impacts the processing of words, but also affects novel words learning both in L1 and L2. For instance, Palmer et al. (2013) asked native English speakers to learn rare English words (novel words) paired with definitions. Half of the novel words were concrete, and the other half abstract. It was found that participants' responses to the concrete novel words were faster than those to the abstract novel words both in semantic categorization task and lexical decision task. In L2 novel word learning, De Groot and Keijzer (2000) used paired-associate training technique, in which a Dutch word and an English pseudoword were visually presented simultaneously. Actually, the pseudowords were letter strings which were orthographically and phonologically legal. Native Dutch speakers performed recall tests immediately following learning and 1 week later. The recall

accuracies for the concrete words were larger than those for the abstract words in both the tests. These results indicate that concrete words are easier to be learned than abstract words both in L1 and L2 vocabulary learning.

Contextual learning, in which people derive meanings of novel words by reading sentences or discourses, is an important approach to word acquisition (Swanborn and de Glopper, 1999; Batterink and Neville, 2011). Previous studies have found that learners can successfully infer the meaning of unknown words rapidly from contexts (Mestres-Missé et al., 2007, 2010; Borovsky et al., 2010, 2012, 2013; Chen et al., 2014; Ding et al., 2017; Zhang et al., 2017). For instance, Mestres-Missé et al. (2007) asked participants to read sentences ending in novel words (pseudowords) or real words. It was found that, only after three exposures in sentences, the N400 amplitudes elicited by the novel words were similar to those elicited by the real words. Using the same learning paradigm, researchers investigated the influence of the concreteness of concepts on novel words learning (Mestres-Missé et al., 2009, 2014). It was found that the meaning of concrete and abstract novel words could be similarly identified. However, the reading times for the abstract novel words on the second sentences were longer than those on the first sentences, and the reading times for the concrete novel words showed the reverse pattern. These results implied that participants had to collect and recheck the information provided by both sentences for abstract new words, indicating that concrete word meaning was discovered and learned faster than abstract word meaning (Mestres-Missé et al., 2014). Furthermore, an fMRI study revealed that learning concrete and abstract novel words was qualitatively different in neural correlates and recruited similar brain regions as the processing of real concrete and abstract words. In particular, the ventral anterior fusiform gyrus, a region driven by imageability (Ishai et al., 2000), was exclusively activated in the association of new concrete words to their meaning, indicating the involvement of nonverbal imagery-based system in concrete novel word leaning (Mestres-Missé et al., 2009).

The above-mentioned studies suggest a learning advantage of concrete novel words over abstract novel words during contextual learning. However, the novel words learned from contexts are not stored in isolation, they can be integrated into semantic memory rapidly (Mestres-Missé et al., 2007; Borovsky et al., 2012; Ding et al., 2017; Zhang et al., 2017). For example, Borovsky et al. (2012) found that the meaning of novel words embedded in highly constraining sentences could be learned based on the high cloze probabilities (mean = 0.896) in the pretest. Furthermore, the novel words could be associated with their semantically related words, as reflected by the reduced N400s compared to the unrelated words in a lexical decision task immediately after learning. Based on the above-mentioned results, the current study aimed to further explore whether there is a concreteness effect in the integration of novel words into semantic memory using event-related potential (ERP) technique. Specifically, we aimed to investigate how the concreteness of concepts influences the learning of novel words and their associations with known words. Most previous studies investigating word acquisition in contextual learning used

pseudowords as new labels for familiar concepts (e.g., Mestres-Missé et al., 2008, 2010; Borovsky et al., 2010; Chen et al., 2014; Ding et al., 2017; Zhang et al., 2017), which can be thought of as simulating second language (L2) learning (Ferreira et al., 2015). In L2 word learning, Dittinger et al. (2016) asked speakers of French to learn Thai words through picture-word associations and found that the meaning of novel words could be learned and associated with semantically related concepts following a short learning. Furthermore, music training could enhance this kind of leaning in adults (Dittinger et al., 2016) and children (Dittinger et al., 2017). The present study paired pseudowords with familiar concepts via contextual learning paradigm and could shed new light on the influence of the concreteness of concepts on L2 novel word learning.

The novel words were Chinese two-character pseudowords and embedded in two-sentence contexts. The corresponding concepts of the novel words were either concrete or abstract. Participants read the contexts and inferred the meaning of the novel words. After learning, they performed a lexical decision task with ERPs being recorded. The learned novel words served as the primes, and the corresponding concepts of the novel words, semantically related words and unrelated words served as the targets. The N400 is a negative-going wave that peaks ~400 ms after the onset of the meaningful stimulus (Kutas and Hillyard, 1980; Kutas and Federmeier, 2011). It is correlated with semantic priming, with the target words eliciting smaller N400 amplitudes when preceded by semantically related words compared to unrelated words in the lexical decision task (Bentin et al., 1985; Estes and Jones, 2009; Borovsky et al., 2012; Jones and Golonka, 2012). Therefore, we expected smaller N400s for the corresponding concepts of the novel words and the semantically related words relative to the unrelated words in both the concrete and abstract conditions. If the concreteness of concepts influences the integration of novel words into semantic memory, the N400 effects would be different between the concrete and abstract conditions.

In addition, a late positive component (LPC) is modulated by the semantic relatedness between the prime and target, with targets preceded by semantically related primes eliciting larger LPCs than those preceded by unrelated primes in the semantic priming lexical decision task (Bouaffre and Faïta-Ainseba, 2007; Kim et al., 2012). This LPC effect reflects conscious awareness of semantic relationship between primes and targets (Bouaffre and Faïta-Ainseba, 2007; Chen et al., 2014). If the concreteness of concepts influences this late processing stage, we expected different LPC effects between the concrete and abstract conditions.

MATERIALS AND METHODS

Participants

Twenty-four university students (mean age 23 years, 12 males) participated in the experiment. They were all right handed native Chinese speakers with normal or corrected-to-normal vision. None of them had dyslexia or neural impairment. This study was approved by the Institutional Review Board of the Institute of Psychology, Chinese Academy of Sciences. Before

the experiment, all subjects read and signed a written informed consent in accordance with the Declaration of Helsinki.

Materials

Sixty-six two-character pseudowords served as the novel words. Half of the corresponding concepts of the novel words are concrete and the other half are abstract. They were embedded in the learning discourses, each of which consisted of two sentences. The first sentences always ended in the novel words. **Table 1** presents the examples of stimuli. We tested the cloze probabilities of the corresponding concepts in the first sentences and the inferring probabilities of the novel words in the whole discourses. In the cloze probability test, participants read the sentences without the final critical words. They were asked to finish the sentences with the first words that came to mind. In the inferring probability test, participants read all the discourses and inferred the meaning of the novel words. Twelve participants firstly took part in the cloze probability test, then in the inferring probability test. The cloze probabilities of the concrete and abstract corresponding concepts were equally high [correct rates: mean (SD) = 80.8% (21.60%) and 81.06% (20.86%) for the concrete and abstract conditions, respectively: $t_{(64)} = -0.10$, $p = 0.924$]. Meanwhile, the concrete and abstract novel words could be successfully inferred at equally high accuracies [correct rates: mean (SD) = 96.97% (4.57%) and 96.72% (5.49%) for the concrete and abstract novel words, respectively: $t_{(64)} = 0.20$, $p = 0.840$].

After the learning phase, the novel words served as primes in a lexical decision task paired with three types of real-word targets and three unlearned-pseudoword targets. The real-word targets included the corresponding concepts (CC), semantically related words (SR), and unrelated words (UR).

TABLE 1 | Examples of the stimuli in the learning phase and lexical decision task.

LEARNING DISCOURSES IN THE LEARNING PHASE	
Concrete condition	**Abstract condition**
小蝌蚪长大之后会变成一只芋沌，此刻池塘里的荷叶上蹲着一只芋沌在捕食昆虫。(A little tadpole grows up into a **yudun**. Right now on a lotus leaf there is a **yudun** catching insects.)	获得诺贝尔奖是科学家的最高栗芸，大家都在为赢得这份最高的栗芸而努力。(To a scientist wining the Nobel prize is the highest **liyun**. Everyone is trying hard to win this highest **liyun**.)

TARGETS IN THE LEXICAL DECISION TASK		
	Concrete condition	**Abstract condition**
Corresponding concept (CC)	青蛙 (frog)	荣誉 (honor)
Semantically related word (SR)	蜥蜴 (lizard)	声望 (reputation)
Unrelated word (UR)	裤子 (pants)	说法 (statement)
Pseudoword	晾炗 (liang ji)	贡颠 (gong dian)
Pseudoword	甚筋 (shen jin)	募旺 (mu wang)
Pseudoword	泉愧 (quan kui)	屑泊 (xie bo)

The examples are presented in Chinese with English translations in parenthesis for the learning discourses and the target words in the lexical decision task. The novel words serving as the primes are in boldface in the discourses.

For the concrete condition, the semantically related words were taxonomically/categorically related to the corresponding concepts. For the abstract condition, the semantically related words were synonyms of the corresponding concepts. Abstract words and their synonyms share semantic similarities, and their semantic relationship is similar to the categorical relationship between concrete words (Crutch and Warrington, 2005; Crutch et al., 2006). The unrelated words were concrete and abstract in the concrete and abstract conditions, respectively. Fourteen participants who did not participate in the cloze probability and inferring probability tests rated the semantic relatedness between the novel words and the semantically related words, as well as between the novel words and the unrelated words on a 7-point Likert scale (7 indicates the most closely related and 1 indicates unrelated). Since the novel words were not known before the experiment, they were replaced by the corresponding concepts. **Table 2** presents the rating results. We conducted a repeated measures ANOVA, with Target condition (SR, UR) serving as a within-item factor and Word category (concrete, abstract) as a between-item factor. There was only a significant main effect of Target condition [$F_{(1, 64)} = 3074.01$, $p < 0.001$, $\eta_p^2 = 0.980$], indicating that the SR targets were more related to the novel words than the UR targets. Neither the main effect of Word category [$F_{(1, 64)} = 2.65$, $p = 0.109$, $\eta_p^2 = 0.040$] nor the interaction between Target condition and Word category [$F_{(1, 64)} = 0.44$, $p = 0.509$, $\eta_p^2 = 0.007$] was significant.

Another 15 participants rated all the target words in concreteness, emotional valence and arousal on 7-point Likert scales (7 indicates the most concrete, most positive, and most aroused). Meanwhile, we checked the word frequency based on the corpus developed by Cai and Brysbaert (2010) and calculated the number of strokes of all the target words. **Table 2** presents the results of the ratings and calculations. We performed separate repeated measures ANOVAs for the concreteness, emotional valence, emotional arousal, word frequency, and number of strokes, with Target condition (CC, SR, UR) serving as a within-item factor and Word category (concrete, abstract) as a between-item factor. **Table 3** presents F-values of the ANOVAs on the stimuli properties. The results showed that the target words differed in concreteness, with the concrete words being more concrete than the abstract words. In addition, all the words were matched for emotional valence and arousal, as well as word frequency and number of strokes.

In addition, to make sure that participants learned the corresponding concepts of the novel words instead of their semantically related words since they first encountered the novel words, we calculated the cloze probabilities in the first sentences and the inferring probabilities in the whole discourses for the semantically related words. The results showed that the cloze probabilities of the concrete and abstract semantically related words were not significantly different [correct rates: mean (SD) = 0.51% (2.02%) and 0.25% (1.45%) for the concrete and abstract semantically related words, respectively: $t_{(64)} = 0.83$, $p = 0.412$], and were not significantly different from zero [$t_{(32)} = 1.36$, $p = 0.184$ and $t_{(32)} = 1.00$, $p = 0.325$ for the concrete and abstract conditions, respectively]. Similarly, their inferring probabilities were equally low [correct rates: mean (SD) = 0.76% (3.20%)

and 0.25% (1.45%) for the concrete and abstract semantically related words, respectively: $t_{(64)} = 0.58$, $p = 0.562$], and were not significantly different from zero [$t_{(32)} = 1.44$, $p = 0.160$ and $t_{(32)} = 1.00$, $p = 0.325$ for the concrete and abstract conditions, respectively].

Procedure

Participants who did not take part in any pretests of the stimuli were seated in a comfortable chair with a distance of about 80 cm from the computer screen. The words were presented in white color on a black screen with a font size of 20 in Song Typeface. Similar to the learning procedure of previous ERP studies (e.g., Zhang et al., 2017), a learning trial started with a 1,000-ms fixation cross in the center of the screen. Then a sentence was presented one word or two-word phrase at a time (500-ms duration, 800-ms stimulus onset asynchrony). The novel words always appeared in isolation for 1,000 ms. After the last phrase, the whole learning discourse was presented on the screen. Participants were asked to press the space button if they had inferred the meaning of the novel word. A 2,000-ms resting screen was presented before the next trial began.

In the lexical decision task, a trial also started with a 1,000-ms fixation cross in the center of the screen. Then a prime word was presented for 300 ms and followed by a 200-ms blank screen. After that, a target word was presented for 300 ms. Participants were asked to judge whether the target was a real word or not as quickly and accurately as possible by pressing the "F" or "J" buttons on the keyboard. The correspondence between F/J and word/pseudoword was counterbalanced across participants.

We divided the 66 discourses into six blocks. In order to balance the number of concrete and abstract items, three blocks included 10 discourses (five concrete and five abstract), and the other three blocks included 12 discourses (six concrete and six abstract). The learning phase and lexical decision task were interleaved. Participants read the discourses in a pseudo-random order in each block with no more than three discourses of the same condition being presented in succession. All word pairs in the lexical decision task were arranged in a random order first. Then, for the word pairs containing the same novel word, the novel word-CC target pair was always presented after the novel word-SR and novel word-UR target pairs. This manipulation was performed to avoid acquisition or recognition of the novel words' meaning through the pairing with their corresponding concepts, which would confound the contextual learning effect. Furthermore, no trial type occurred more than three times consecutively and trials containing the same novel words were not presented in succession. Finally, the word pairs were presented in a pseudo-random order. There was a short break between blocks.

Electrophysiological Recording and Preprocessing

EEG was recorded with 64 Ag/AgCl electrodes mounted on an elastic cap at a sampling rate of 500 Hz with a band pass filter of 0.05–100 Hz. EEG data were amplified with AC amplifiers. The right mastoid electrode served as the online reference, and an electrode placed between Fz and FPz electrodes served as

TABLE 2 | Means (SDs) of the stimuli properties.

	Corresponding concepts		Semantically related words		Unrelated words	
	Concrete	Abstract	Concrete	Abstract	Concrete	Abstract
Relatedness	–	–	5.21 (0.44)	5.26 (0.51)	1.39 (0.20)	1.59 (0.27)
Concreteness	6.48 (0.39)	2.91 (0.61)	6.41 (0.48)	2.87 (0.59)	6.41 (0.28)	2.77 (0.60)
Valence	4.68 (0.73)	4.63 (0.95)	4.51 (0.83)	4.61 (0.85)	4.75 (0.70)	4.36 (0.63)
Arousal	2.80 (0.79)	3.03 (0.84)	2.61 (0.66)	2.85 (0.75)	2.62 (0.74)	2.67 (0.74)
Word frequency	2.57 (0.97)	2.52 (0.89)	2.50 (0.94)	2.36 (0.99)	2.51 (0.73)	2.34 (0.86)
Number of strokes	15.61 (4.77)	16.72 (5.10)	16.36 (4.74)	16.03 (5.14)	17.00 (3.62)	16.09 (4.08)

TABLE 3 | F-values of the ANOVAs on the stimuli properties.

	Concreteness	Valence	Arousal	Word frequency	Number of strokes
Target condition	0.76	0.40	2.54	0.48	0.14
Word category	2640.97***	0.73	2.06	0.62	0.95
Target condition by Word category	0.15	2.12	0.40	0.11	0.88

*The df for Target condition and the interaction was (2, 128), for Word category was (1, 64). *** Significant at 0.001 level.*

the ground. An electrode was also placed over the left mastoid. Two electrodes above and below the left eye were used to monitor the vertical eye movements and blinks. The horizontal eye movements were monitored via two electrodes placed lateral to the outer canthus of each eye. Impedance of most electrodes was kept below 5 kΩ.

The raw EEG data were preprocessed with NeuroScan software 4.5 offline. After automatic correction of the ocular artifacts (Semlitsch et al., 1986), the EEG data were filtered using a band-pass filter at 0.1–30 Hz and segmented into 1,200-ms epochs from −200 to 1,000 ms relative to the target words onset. The mean amplitudes in the prestimulus interval served as baseline. An artifact correction of ±80 µV was used at all electrodes except the electrooculograms. Then, the ERPs were re-referenced offline to the average of two mastoids. Finally, average ERPs were calculated for each participant at each electrode in each condition.

ERP Data Analysis

The mean amplitudes calculated for each participant, each condition, within each selected time window were entered into statistical analysis. **Figure 1** shows the selected electrodes for analysis. Target condition (CC, SR, UR), Word category (concrete, abstract), Laterality (left, middle, right), and Anteriority (anterior, central, posterior) were taken as within-subject factors in repeated measures ANOVAs. In addition, simple effect tests and planned comparisons were conducted when there were any interactions with the critical manipulations in ANOVAs. Bonferroni correction was applied to adjust the multiple comparisons. The original degrees of freedom were reported with corrected p-values according to
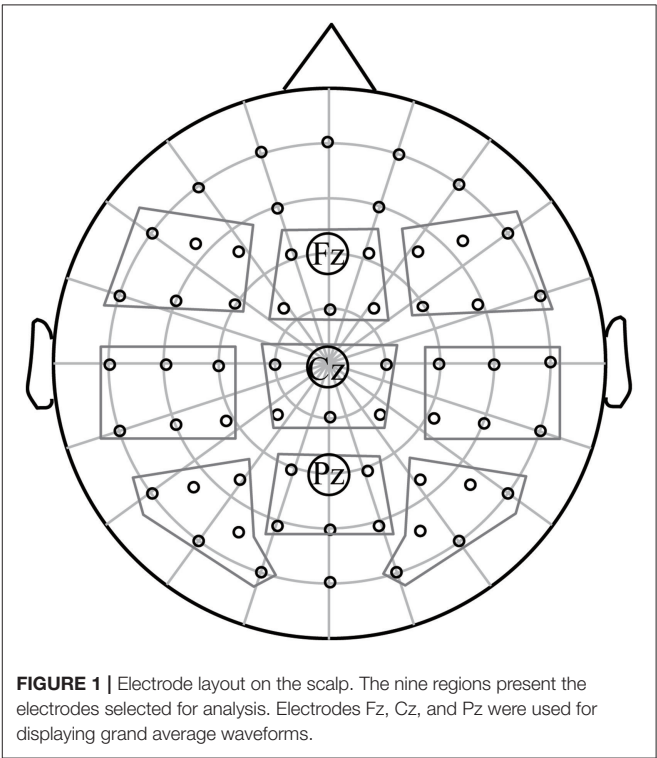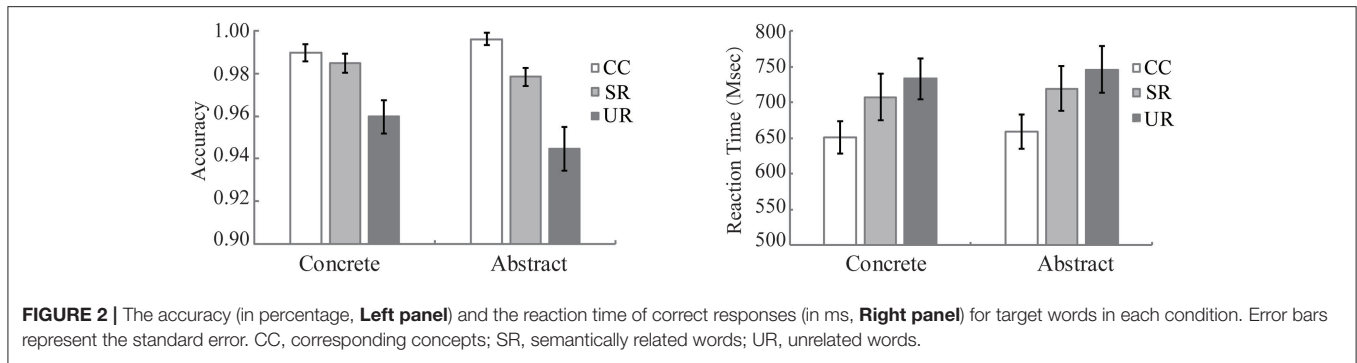

FIGURE 1 | Electrode layout on the scalp. The nine regions present the electrodes selected for analysis. Electrodes Fz, Cz, and Pz were used for displaying grand average waveforms.

the Greenhouse-Geisser correction applied when appropriate (Greenhouse and Geisser, 1959).

RESULTS

Behavioral Data

Figure 2 presents the accuracy results (left panel) and the reaction time results (right panel). We conducted 3 Target condition (CC, SR, UR) by 2 Word category (concrete, abstract) repeated measures ANOVAs for the accuracy and reaction time data. For accuracy, there was a significant main effect of Target condition [$F_{(2, 46)} = 20.76$, $p < 0.001$, $\eta_p^2 = 0.474$]. Pair-wise comparisons revealed that the accuracy for the CC targets was the highest [CC vs. SR: $t_{(23)} = 3.67$, $p = 0.007$; CC vs. UR: $t_{(23)} = 5.86$, $p < 0.001$]. In addition, the accuracy for the SR targets was higher than that for the UR targets [SR vs. UR: $t_{(23)} = 3.75$, $p = 0.009$]. The main effect of Word category [$F_{(1, 23)} = 2.77$, $p =$

FIGURE 2 | The accuracy (in percentage, **Left panel**) and the reaction time of correct responses (in ms, **Right panel**) for target words in each condition. Error bars represent the standard error. CC, corresponding concepts; SR, semantically related words; UR, unrelated words.

0.110, $\eta_p^2 = 0.107$] or the interaction between Word category and Target condition [$F_{(2, 46)} = 2.03$, $p = 0.161$, $\eta_p^2 = 0.081$] was not significant.

For reaction time, error trials and outlier data points which were 2.5 standard deviations away from the mean were excluded from analysis. The repeated ANOVA revealed a significant main effect of Target condition [$F_{(2, 46)} = 20.08$, $p < 0.001$, $\eta_p^2 = 0.466$]. Pair-wise comparisons revealed that the participants responded faster for the CC targets than for the SR [CC vs. SR: $t_{(23)} = -3.80$, $p = 0.003$] and UR [CC vs. UR: $t_{(23)} = -5.64$, $p < 0.001$] targets. In addition, the responses to the SR targets were faster than those for the UR targets [SR vs. UR: $t_{(23)} = -2.63$, $p = 0.045$]. Neither the main effect of Word category [$F_{(1, 23)} = 2.41$, $p = 0.134$, $\eta_p^2 = 0.095$] nor the interaction between the two factors [$F_{(2, 46)} = 0.15$, $p = 0.861$, $\eta_p^2 = 0.006$] was significant.

ERP Data

The grand average waveforms elicited by the target words in the concrete (Left panel) and abstract (Middle panel) conditions, as well as difference waveforms between the concrete and abstract conditions (Right panel) were presented at Fz, Cz, and Pz electrodes in **Figure 3**. Based on visual inspection and previous studies (e.g., Zhang et al., 2017), two time windows were selected for statistical analysis: (1) the standard N400 time window: 300–500 ms; (2) the LPC time window: 600–800 ms.

The statistical analysis of the N400 component revealed a significant main effect of Target condition [$F_{(2, 46)} = 15.88$, $p < 0.001$, $\eta_p^2 = 0.408$]. Pair-wise comparisons showed that the CC targets elicited the smallest N400 amplitudes [CC vs. SR: $t_{(23)} = 3.09$, $p = 0.015$; CC vs. UR: $t_{(23)} = 5.27$, $p < 0.001$]. Meanwhile, the SR targets elicited smaller N400s than the UR targets [SR vs. UR: $t_{(23)} = 2.71$, $p = 0.037$]. The N400 effects were distributed over all electrodes tested. There were not any other significant effects.

In the 600–800 ms time window, there was a significant interaction between Target condition and Word category [$F_{(2, 46)} = 3.78$, $p = 0.030$, $\eta_p^2 = 0.141$]. Simple effect tests revealed a significant effect of Target condition in the concrete condition [$F_{(2, 46)} = 4.30$, $p = 0.019$, $\eta_p^2 = 0.157$], but not in the abstract condition [$F_{(2, 46)} = 0.90$, $p = 0.413$, $\eta_p^2 = 0.038$]. Pair-wise comparisons showed that the concrete CC targets elicited a

smaller LPC than the concrete UR targets [CC vs. UR: $t_{(23)} = -3.59$, $p = 0.005$; CC vs. SR: $t_{(23)} = -2.22$, $p = 0.109$; SR vs. UR: $t_{(23)} < 1$] over all electrodes tested. No other significant main effects or interactions of interest were observed.

DISCUSSION

The current study aimed to examine whether and how the concreteness of concepts influences the learning of novel words and their associations with known words in semantic memory. After inferring the meaning of the novel words in the learning contexts, participants completed a lexical decision task with the learned novel words serving as the prime words. The corresponding concepts of the novel words and the semantically related words were judged faster and more accurately, and elicited smaller N400s compared to the unrelated words. The N400 effect was not modulated by the concreteness of concepts. In addition, the corresponding concepts of the concrete novel words elicited smaller LPCs than the concrete unrelated words. This LPC effect was absent in the abstract condition.

The learned novel words, irrespective of the concreteness of concepts, facilitated the processing of their corresponding concepts and semantically related words, as reflected by the behavioral data and the N400 effects. This is in line with previous studies investigating novel word acquisition in contextual learning (Mestres-Missé et al., 2007; Borovsky et al., 2012; Chen et al., 2014; Ding et al., 2017; Zhang et al., 2017). These results indicate that learners can successfully infer the meaning of novel words from highly constraining contexts and rapidly connect the novel words with their semantically related words in semantic memory.

Unlike the N400 effect, the LPC effect was modulated by the concreteness of concepts in the current study. The LPC effect has been proposed to reflect conscious awareness of semantic relationship between the prime and the target at a late processing stage (Bouaffre and Faïta-Ainseba, 2007; Chen et al., 2014), with related targets eliciting larger late positive waveforms than unrelated targets (Brown et al., 2000; Hill et al., 2002). However, the LPC effect obtained in the current study might not reflect the semantic relationship-detection (Hill et al., 2005) for two reasons. First, the LPC effect was observed for the corresponding concepts of the novel words, not for the
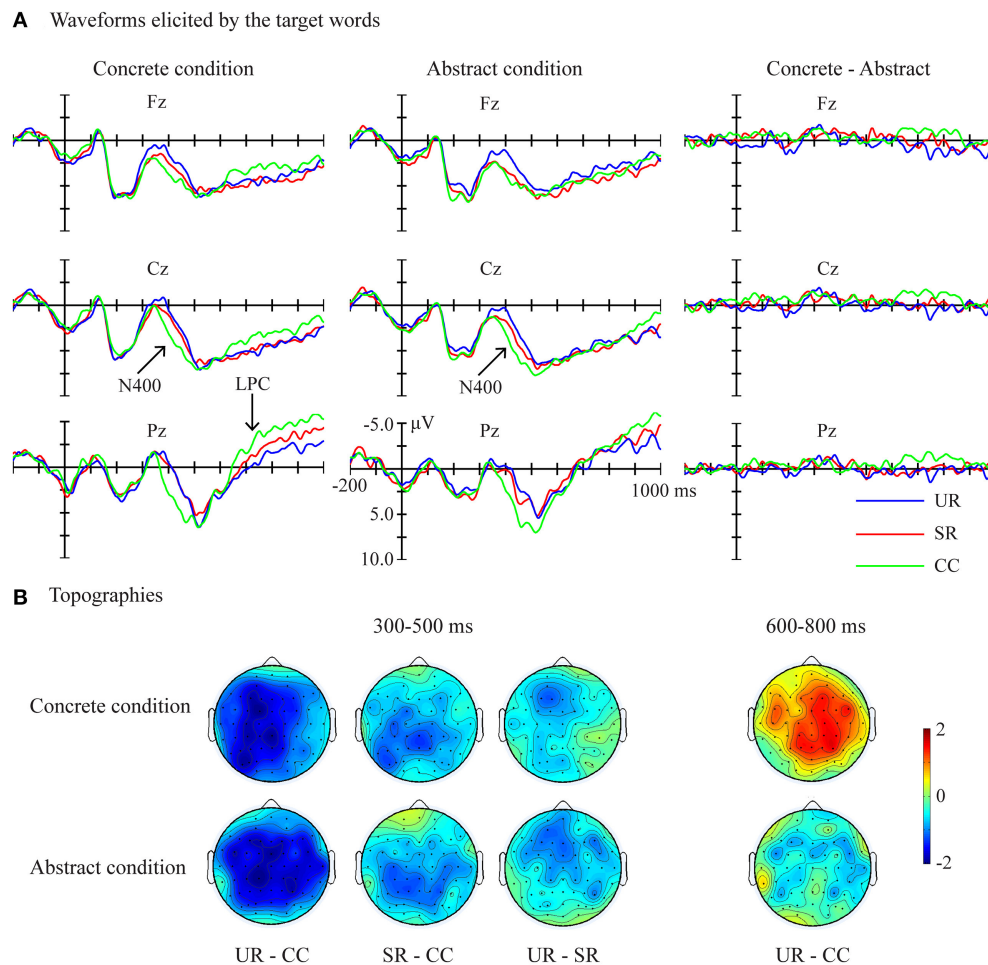
FIGURE 3 | Results of the ERP analysis. (A) Waveforms elicited by the CC, SR, and UR targets in the concrete (Left panel) and abstract (Middle panel) conditions, as well as difference waveforms between the concrete and abstract conditions (Right panel) were presented at Fz, Cz, and Pz electrodes. **(B)** Topographies showing the average amplitude voltage differences between the CC, SR, and UR targets, respectively, in the time windows of 300–500 and 600–800 ms. CC, corresponding concepts; SR, semantically related words; UR, unrelated words.

semantically related words. Second, the corresponding concepts elicited smaller, but not larger, LPCs than the unrelated words. We propose two possibilities for the LPC effect as follows. Firstly, the LPC effect is correlated with processing difficulty (Brouwer et al., 2012); therefore, the observed LPC reduction reflected a facilitation of the concrete novel words on the processing of their corresponding concepts. In the present study, the novel words were new forms of the corresponding concepts. After learning, the novel words were stored as new labels of their corresponding concepts in semantic memory. Furthermore, due to the processing advantage of the concrete concepts, the association strength between the novel words and the concepts might be stronger for the concrete words than for the abstract words. Hence, the concrete novel words were learned more deeply than the abstract novel words in contextual reading. Secondly, the LPC has been viewed as part of the P300 family of ERP components that reflect context updating (Donchin

and Coles, 1988). The larger LPCs to the unrelated words than to the corresponding concepts of the novel words in the concrete condition possibly reflected more context updating for the unrelated words, because they were less expected following the primes than the corresponding concepts. In other words, the concrete learning contexts were recollected more vividly than the abstract learning contexts and potentially more so for the concrete novel words. The process may facilitate the processing of their corresponding concepts relative to the unrelated words. This interpretation implies that episodic memory played an important role in the testing phase. It should be noted that the second possibility is not contradictory to the integration of novel words into semantic memory, as indicated by the N400 effect given its automaticity (Kutas and Federmeier, 2011).

According to the context availability hypothesis, the concreteness effect would disappear with the provision of contexts to concrete and abstract words (Schwanenflugel and

Shoben, 1983; Schwanenflugel et al., 1988). For instance, the lexical decision times for the concrete and abstract words were equivalent, with a sentence providing contextual information (Schwanenflugel and Shoben, 1983). The behavioral data did not reveal a difference between the concrete and abstract conditions, which is consistent with the context availability theory. However, the context availability theory could not explain the LPC effect we observed in the concrete condition. This concreteness effect could be accounted for by the dual coding theory (Paivio, 1986) or the extended dual coding theory (Holcomb et al., 1999). The usage of the imagery-based system or more verbal information, or both of them, for the concrete concepts facilitated the learning of the concrete novel words. All these results suggest that concrete and abstract words might be quantitatively and qualitatively different. We propose, as previous studies suggested, concreteness of concepts is not a dual feature, but a continuum (e.g., Mestres-Missé et al., 2014).

It is worth to note that we also did not find a concreteness effect (i.e., main effect of Word category) in the N400 time window. Previous studied have found that concrete words elicit larger N400s than abstract words (e.g., Kounios and Holcomb, 1994; Zhang et al., 2006; Tolentino and Tokowicz, 2009; Tsai et al., 2009), reflecting more activation of semantic information from memory for concrete words (Kounios and Holcomb, 1994). This N400 effect in response to the concreteness of concepts was also observed in novel word learning studies (e.g., Palmer et al., 2013). The absence of the concreteness effect in the N400 amplitudes could be attributed to the similar context availabilities for both the concrete and abstract novel words as discussed above. However, one might argue that the concrete and abstract unrelated words might differ in context availability because they did not appear in the learning contexts. Thus, there should be a processing advantage for the unrelated words in the concrete condition over the abstract condition. We propose that the absence of the concreteness effect might be alternatively due to the relative high word frequency of the target words. Previous behavioral researches revealed no concreteness effects for high frequency words (e.g., James, 1975; De Groot, 1989; Miller and Roodenrys, 2009). The average word frequency of the critical words in the current study was 2.47, which is the log transform of the total number of times that the word appears in film subtitles (mean = 0.85, median = 0.60, mode = 0; Cai and Brysbaert, 2010). This relatively high word frequency might lead to the disappearance of the concreteness effect.

However, Zhang et al. (2006) found that concrete nouns elicited larger N400s than abstract nouns, regardless of word frequency, indicating that the concreteness effect is immune to word frequency. The different results between the current study and the study of Zhang et al. (2006) might have resulted from the discrepancies in the experimental procedure. Zhang et al. only asked participants to perform a lexical decision task, while in the current study, participants performed a lexical decision task in the semantic priming paradigm with novel words serving as primes. Hill et al. (2005) found that the N400 component was larger in the long SOA (700 ms) than in the short SOA (150 ms) for the real word targets when pseudowords served as

primes. They attributed this N400 enhancement to the use of pseudoword primes which may drive the subjects to a deeper semantic processing of both the primes and targets. Therefore, the concreteness effects reflected in the N400 component might be superimposed on the deeper semantic processing. In a word, the provision of contexts and relatively high word frequency, as well as the experimental procedure might result in the absence of the concreteness effect in the behavioral data and the N400 amplitudes.

In addition, there were graded increasing N400 amplitudes for the corresponding concepts, the semantically related words and the unrelated words. These N400 priming effects were not modulated by the concreteness of concepts. First, as discussed above, these results suggest that novel words could be integrated into semantic memory rapidly, which is consistent with previous studies using contextual learning (Borovsky et al., 2012; Ding et al., 2017; Zhang et al., 2017) and picture-word associations (Dittinger et al., 2016). Second, these results are in conflict with the structurally different representational frameworks theory (Crutch and Warrington, 2005), which also proposes that the representations of concrete and abstract words differ qualitatively. Crutch and Warrington (2005) asked a patient with semantic refractory access dysphasia to perform a spoken word-written word matching task, in which the patient was required to point to the target written word in a word array following the spoken word. Words in the same array related to each other via either semantic association or semantic similarity. The results revealed interference for the semantically associated abstract words, but not for the semantically synonymous abstract words. However, the concrete words showed the reverse pattern. Based on these results, researchers proposed that the concrete concepts are organized by semantic similarity, and abstract concepts are represented in an associated neural network (Crutch and Warrington, 2005; Crutch et al., 2006). However, subsequent studies on healthy participants (Zhang et al., 2013; Geng and Schnur, 2015) and patients (Hamilton and Coslett, 2008) found that semantic similarity is also important to abstract concepts. In the current study, when preceded by the learned novel words, the synonyms of the abstract words serving as the semantically related words elicited smaller N400s than the unrelated words, indicating that the semantic similarity between concepts also plays a role in the representations of abstract words. These results again support the quantitative instead of the qualitative differences between the concrete and abstract words.

There are some limitations to this study. Firstly, the novel word-CC target pairs were always presented after the novel word-SR target and the novel word-UR target pairs. This design was to avoid the meaning of the novel words being acquired by the pairing of the novel words with their corresponding concepts. Although the design feature was equally true for the abstract and concrete words, it somewhat complicated the interpretation of the results. In future studies, more experimental stimuli and a Latin-square design for the three types of the prime-target pairs would address this issue. Secondly, the difference in the vividness between the concrete and abstract contexts, as well as the immediate test following the learning

phase made it possible that episodic memory played an important role in semantic processing. Testing the integration of novel words into the semantic network at least a day after learning may partially rule out the contribution of episodic memory.

In summary, both the concrete and abstract novel words learned from contextual information could prime their corresponding concepts and the semantically related words compared to the unrelated words, as reflected by the graded increasing N400s for the three types of words. In addition, the concrete novel words impacted the processing of their corresponding concepts at a late processing stage as indicated by the LPC effect in the concrete condition. This study demonstrated that learners can infer the meaning of concrete and abstract novel words from contextual information, and integrate them into the semantic network. Furthermore, the concrete novel words are learned better than the abstract novel words. Because the word learning task in this experiment resembles second language learning, findings from this investigation shed new light on adult word learning and changes in concept organization.

AUTHOR CONTRIBUTIONS

JD and YY: conceived the idea of the study; JD and WL: collected and analyzed the data; All authors contributed to the interpretation, drafting, critical revision, and final approval of the manuscript for publication.

FUNDING

REFERENCES

Batterink, L., and Neville, H. (2011). Implicit and explicit mechanisms of word learning in a narrative context: an event-related potential study. *J. Cogn. Neurosci.* 23, 3181–3196. doi: 10.1162/jocn_a_00013

Bentin, S., McCarthy, G., and Wood, C. C. (1985). Event-related potentials, lexical decision and semantic priming. *Electroencephalogr. Clin. Neurophysiol.* 60, 343–355. doi: 10.1016/0013-4694(85)90008-2

Binder, J. R., Westbury, C. F., McKiernan, K. A., Possing, E. T., and Medler, D. A. (2005). Distinct brain systems for processing concrete and abstract concepts. *J. Cogn. Neurosci.* 17, 905–917. doi: 10.1162/0898929054021102

Borovsky, A., Elman, J. L., and Kutas, M. (2012). Once is enough: N400 indexes semantic integration of novel word meanings from a single exposure in context. *Lang. Learn. Dev.* 8, 278–302. doi: 10.1080/15475441.2011.614893

Borovsky, A., Kutas, M., and Elman, J. L. (2010). Learning to use words: event-related potentials index single-shot contextual word learning. *Cognition* 116, 289–296. doi: 10.1016/j.cognition.2010.05.004

Borovsky, A., Kutas, M., and Elman, J., (2013). Getting it right: word learning across the hemispheres. *Neuropsychologia* 51, 825–837. doi: 10.1016/j.neuropsychologia.2013.01.027

Bouaffre, S., and Faïta-Ainseba, F. (2007). Hemispheric differences in the time-course of semantic priming processes: evidence from event-related potentials (ERPs). *Brain Cogn.* 63, 123–135. doi: 10.1016/j.bandc.2006.10.006

Brouwer, H., Fitz, H., and Hoeks, J. (2012). Getting real about semantic illusions: rethinking the functional role of the P600 in language comprehension. *Brain Res.* 1446, 127–143. doi: 10.1016/j.brainres.2012.01.055

Brown, C. M., Hagoort, P., and Chwilla, D. J. (2000). An event-related brain potential analysis of visual word priming effects. *Brain Lang.* 72, 158–190. doi: 10.1006/brln.1999.2284

Cai, Q., and Brysbaert, M. (2010). SUBTLEX-CH: chinese word and character frequencies based on film subtitles. *PLoS ONE* 5:e10729. doi: 10.1371/journal.pone.0010729

Chen, S., Wang, L., and Yang, Y. (2014). Acquiring concepts and features of novel words by two types of learning: direct mapping and inference. *Neuropsychologia* 56, 204–218. doi: 10.1016/j.neuropsychologia.2014.01.012

Crutch, S. J., Ridha, B. H., and Warrington, E. K. (2006). The different frameworks underlying abstract and concrete knowledge: evidence from a bilingual patient with a semantic refractory access dysphasia. *Neurocase* 12, 151–163. doi: 10.1080/13554790600598832

Crutch, S. J., and Warrington, E. K. (2005). Abstract and concrete concepts have structurally different representational frameworks. *Brain* 128, 615–627. doi: 10.1093/brain/awh349

De Groot, A. M. (1989). Representational aspects of word imageability and word frequency as assessed through word association. *J. Exp. Psychol. Learn. Mem. Cogn.* 15, 824–845. doi: 10.1037/0278-7393.15.5.824

De Groot, A. M., and Keijzer, R. (2000). What is hard to learn is easy to forget: the roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Lang. Learn.* 50, 1–56. doi: 10.1111/0023-8333.00110

Ding, J., Chen, S., Wang, L., and Yang, Y. (2017). Thematic and taxonomic relations of novel words learned from action and perceptual features. *J. Neurolinguistics* 41, 70–84. doi: 10.1016/j.jneuroling.2016.10.002

Dittinger, E., Barbaroux, M., D'Imperio, M., Jäncke, L., Elmer, S., and Besson, M. (2016). Professional music training and novel word learning: from faster semantic encoding to longer-lasting word representations. *J. Cogn. Neurosci.* 28, 1584–1602. doi: 10.1162/jocn_a_00997

Dittinger, E., Chobert, J., Ziegler, J. C., and Besson, M. (2017). Fast brain plasticity during word learning in musically-trained children. *Front. Hum. Neurosci.* 11:233. doi: 10.3389/fnhum.2017.00233

Donchin, E., and Coles, M. G. (1988). Is the P300 component a manifestation of context updating? *Behav. Brain Sci.* 11, 357–374. doi: 10.1017/S0140525X00058027

Estes, Z., and Jones, L. L. (2009). Integrative priming occurs rapidly and uncontrollably during lexical processing. *J. Exp. Psychol. Gen.* 138, 112–130. doi: 10.1037/a0014677

Ferreira, R. A., Göbel, S. M., Hymers, M., and Ellis, A. W. (2015). The neural correlates of semantic richness: evidence from an fMRI study of word learning. *Brain Lang.* 143, 69–80. doi: 10.1016/j.bandl.2015.02.005

Geng, J., and Schnur, T. T. (2015). The representation of concrete and abstract concepts: categorical versus associative relationships. *J. Exp. Psychol. Learn. Mem. Cogn.* 41, 22–41. doi: 10.1037/a0037430

Greenhouse, S. W., and Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika* 24, 95–112. doi: 10.1007/BF02289823

Hamilton, A. C., and Coslett, H. B. (2008). Refractory access disorders and the organization of concrete and abstract semantics: do they differ? *Neurocase* 14, 131–140. doi: 10.1080/13554790802032218

Hill, H., Ott, F., and Weisbrod, M. (2005). SOA-dependent N400 and P300 semantic priming effects using pseudoword primes and a delayed lexical decision. *Int. J. Psychophysiol.* 56, 209–221. doi: 10.1016/j.ijpsycho.2004.12.004

Hill, H., Strube, M., Roesch-Ely, D., and Weisbrod, M. (2002). Automatic vs. controlled processes in semantic priming - differentiation

by event-related potentials. *Int. J. Psychophysiol.* 44, 197–218. doi: 10.1016/S0167-8760(01)00202-1

Holcomb, P. J., Kounios, J., Anderson, J. E., and West, W. C. (1999). Dual-coding, context-availability, and concreteness effects in sentence comprehension: an electrophysiological investigation. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 721–742. doi: 10.1037/0278-7393.25.3.721

Ishai, A., Ungerleider L. G., and Haxby, J. V. (2000). Distributed neural systems for the generation of visual images. *Neuron* 28, 979–990. doi: 10.1016/S0896-6273(00)00168-9

James, C. T. (1975). The role of semantic information in lexical decisions. *J. Exp. Psychol. Hum. Percept. Perform.* 1, 130–136. doi: 10.1037/0096-1523.1.2.130

Jessen, F., Heun, R., Erb, M., Granath, D. O., Klose, U., Papassotiropoulos, A., et al. (2000). The concreteness effect: evidence for dual coding and context availability. *Brain Lang.* 74, 103–112. doi: 10.1006/brln.2000.2340

Jones, L. L., and Golonka, S. (2012). Different influences on lexical priming for integrative, thematic, and taxonomic relations. *Front. Hum. Neurosci.* 6:205. doi: 10.3389/fnhum.2012.00205

Kim, A. S., Binns, M. A., and Alain, C. (2012). Neuroelectric evidence for cognitive association formation: an event-related potential investigation. *PLoS ONE* 7:e34856. doi: 10.1371/journal.pone.0034856

Kounios, J., and Holcomb, P. J. (1994). Concreteness effects in semantic processing: ERP evidence supporting dual-coding theory. *J. Exp. Psychol. Learn. Mem. Cogn.* 20, 804–823. doi: 10.1037/0278-7393.20.4.804

Kutas, M., and Federmeier, K. D. (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

Kutas, M., and Hillyard, S. A. (1980). Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biol. Psychol.* 11, 99–116. doi: 10.1016/0301-0511(80)90046-0

Mestres-Missé, A., Càmara, E., Rodriguez-Fornells, A., Rotte, M., and Münte, T. F. (2008). Functional neuroanatomy of meaning acquisition from context. *J. Cogn. Neurosci.* 20, 2153–2166. doi: 10.1162/jocn.2008.20150

Mestres-Missé, A., Münte, T. F., and Rodriguez-Fornells, A. (2009). Functional neuroanatomy of contextual acquisition of concrete and abstract words. *J. Cogn. Neurosci.* 21, 2154–2171. doi: 10.1162/jocn.2008.21171

Mestres-Missé, A., Münte, T. F., and Rodriguez-Fornells, A. (2014). Mapping concrete and abstract meanings to new words using verbal contexts. *Second Lang. Res.* 30, 191–223. doi: 10.1177/0267658313512668

Mestres-Missé, A., Rodriguez-Fornells, A., and Münte, T. F. (2007). Watching the brain during meaning acquisition. *Cereb. Cortex* 17, 1858–1866. doi: 10.1093/cercor/bhl094

Mestres-Missé, A., Rodriguez-Fornells, A., and Münte, T. F. (2010). Neural differences in the mapping of verb and noun concepts onto novel words. *Neuroimage* 49, 2826–2835. doi: 10.1016/j.neuroimage.2009.10.018

Miller, L. M., and Roodenrys, S. (2009). The interaction of word frequency and concreteness in immediate serial recall. *Mem. Cogn.* 37, 850–865. doi: 10.3758/MC.37.6.850

Paivio, A. (1986). *Mental Representations: A Dual Coding Approach*. New York, NY: Oxford University Press.

Paivio, A. (1991). Dual coding theory: retrospect and current status. *Can. J. Psychol.* 45, 255–287. doi: 10.1037/h0084295

Palmer, S. D., Macgregor, L. J., and Havelka, J. (2013). Concreteness effects in single-meaning, multi-meaning and newly acquired words. *Brain Res.* 1538, 135–150. doi: 10.1016/j.brainres.2013.09.015

Schwanenflugel, P. J., Akin, C., and Luh, W. M. (1992). Context availability and the recall of abstract and concrete words. *Mem. Cogn.* 20, 96–104. doi: 10.3758/BF03208259

Schwanenflugel, P. J., Harnishfeger, K. K., and Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words. *J. Mem. Lang.* 27, 499–520. doi: 10.1016/0749-596X(88)90022-8

Schwanenflugel, P. J., and Shoben, E. J. (1983). Differential context effects in the comprehension of abstract and concrete verbal materials. *J. Exp. Psychol. Learn. Mem. Cogn.* 9, 82–102. doi: 10.1037/0278-7393.9.1.82

Semlitsch, H. V., Anderer, P., Schuster, P., and Presslich, O. (1986). A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP. *Psychophysiology* 23, 695–703. doi: 10.1111/j.1469-8986.1986.tb00696.x

Swanborn, M. S. L., and de Glopper, K. (1999). Incidental word learning while reading: a meta-analysis. *Rev. Educ. Res.* 69, 261–285. doi: 10.3102/00346543069003261

Tolentino, L. C., and Tokowicz, N. (2009). Are pumpkins better than heaven? An ERP investigation of order effects in the concrete-word advantage. *Brain Lang.* 110, 12–22. doi: 10.1016/j.bandl.2009.01.001

Tsai, P., Yu, B. H., Lee, C. Y., Tzeng, O. J., Hung, D. L., and Wu, D. H., (2009). An event-related potential study of the concreteness effect between Chinese nouns and verbs. *Brain Res.* 1253, 149–160. doi: 10.1016/j.brainres.2008.10.080

West, W. C., and Holcomb, P. J. (2000). Imaginal, semantic, and surface-level processing of concrete and abstract words: an electrophysiological investigation. *J. Cogn. Neurosci.* 12, 1024–1037. doi: 10.1162/08989290051137558

Zhang, M., Chen, S., Wang, L., Yang, X., and Yang, Y. (2017). Episodic specificity in acquiring thematic knowledge of novel words from descriptive episodes. *Front. Psychol.* 8:488. doi: 10.3389/fpsyg.2017.00488

Zhang, Q., Guo, C. Y., Ding, J. H., and Wang, Z. Y., (2006). Concreteness effects in the processing of Chinese words. *Brain Lang.* 96, 59–68. doi: 10.1016/j.bandl.2005.04.004

Zhang, X., Han, Z., and Bi, Y. (2013). Are abstract and concrete concepts organized differently? Evidence from the blocked translation paradigm. *Appl. Psycholinguist.* 34, 1059–1092. doi: 10.1017/S0142716412000124

When the Second Language Takes the Lead: Neurocognitive Processing Changes in the First Language of Adult Attriters

Kristina Kasparian[1,2]* and Karsten Steinhauer[1,2]*

[1] Neurocognition of Language Laboratory, School of Communication Sciences and Disorders, McGill University, Montreal, QC, Canada, [2] Centre for Research on Brain, Language and Music, Montreal, QC, Canada

Although research on multilingualism has revealed continued neuroplasticity for language-learning beyond what was previously expected, it remains controversial whether and to what extent a second language (L2) acquired in adulthood may induce changes in the neurocognitive processing of a first language (L1). First language (L1) attrition in adulthood offers new insight on neuroplasticity and the factors that modulate neurocognitive responses to language. To date, investigations of the neurocognitive correlates of L1 attrition and of factors influencing these mechanisms are still scarce. Moreover, most event-related-potential (ERP) studies of second language processing have focused on L1 influence on the L2, while cross-linguistic influence in the reverse direction has been underexplored. Using ERPs, we examined the real-time processing of Italian relative-clauses in 24 Italian-English adult migrants with predominant use of English since immigration and reporting attrition of their native-Italian (Attriters), compared to 30 non-attriting monolinguals in Italy (Controls). Our results showed that Attriters differed from Controls in their acceptability judgment ratings and ERP responses when relative clause constructions were ungrammatical in English, though grammatical in Italian. Controls' ERP responses to unpreferred sentence constructions were consistent with garden path effects typically observed in the literature for these complex sentences. In contrast, due to L2-English influence, Attriters were less sensitive to semantic cues than to word-order preferences, and processed permissible Italian sentences as outright morphosyntactic violations. Key factors modulating processing differences within Attriters were the degree of maintained L1 exposure, length of residence in the L2 environment and L2 proficiency – with higher levels of L2 immersion and proficiency associated with increased L2 influence on the L1. To our knowledge, this is the first demonstration that high levels of L2 proficiency and exposure may render a grammatical sentence in one's native language ungrammatical. These group differences strongly point to distinct processing strategies and provide evidence that even a "stabilized" L1 grammar is subject to change after a prolonged period of L2 immersion and reduced L1 use, especially in linguistic areas promoting cross-linguistic influence.

Keywords: neuroplasticity, first language attrition, second language acquisition, event-related potentials, language processing, crosslinguistic influence, relative clauses, language exposure

INTRODUCTION

First Language (L1) Attrition

First language (L1) attrition allows us to study the impact of the second language (L2) on the native-language in a context of prolonged L2 immersion and reduced L1 use, usually after immigration to a new country (Köpke and Schmid, 2004). A number of behavioral studies have shown that attrition is typically detectable in the domain of lexical-semantics (de Bot, 1996; Köpke, 1999; Hulsen, 2000; Paradis, 2003, 2007; Köpke and Schmid, 2004; Montrul, 2008; Opitz, 2011), whereas findings have been mixed in the domain of morphosyntax (Ammerlaan, 1996; Gürel, 2004; Tsimpli et al., 2004; Kim et al., 2010; Schmid, 2010; Schmid and Köpke, 2011; Sorace, 2011). Moreover, it has been shown that L1 attrition is far less pervasive in adults than in children (see reviews by Köpke, 2004 and Köpke and Schmid, 2004), in whom L1 linguistic patterns are argued to be deeply entrenched (Tokowicz and MacWhinney, 2005) and stabilized (Bylund, 2009).

While some behavioral studies have provided evidence of L2 influence on the grammar of adult L1 attriters (see Schmid, 2011), neurocognitive investigations of L1 attrition are still scarce (Pallier, 2007; Datta, 2010; Schmid, 2013; Kasparian, 2015). A recent event-related-potential (ERP) study by Bergmann et al. (2015) tested German–English attriters' processing of gender agreement violations and verb form combinations, compared to monolingual native speakers of German. Both groups showed the same ERP response (a posterior P600 effect) when processing gender agreement violations. However, when processing verb form violations, only the attriters showed an additional N400 effect prior to the posterior P600, suggestive of potential influence from their L2-English grammar, in which verb form violations have been found to elicit such biphasic N400+P600 responses (Sabourin and Stowe, 2008). The authors did not report whether these response patterns were modulated by any factors related to the attriters' bilingual experience, such as L1 proficiency, exposure/use, length of residence (LoR), etc. Attriters scored lower than native-monolinguals on a written proficiency measure (German *C*-test, Schmid and Dusseldorp, 2010) but did not differ in their acceptability ratings nor on an offline gender assignment task. The authors concluded that the predominantly used L2 engenders little change to the processing of the deeply entrenched L1 grammar, and that ERPs are less susceptible to attrition effects than active language production.

The opposite was found in a recent ERP study of number agreement processing in Italian by Kasparian et al. (2016). Although attriters (L1-Italian, L2-English) scored numerically lower on a number of written and oral proficiency measures, the only behavioral difference from native-controls that reached significance was the attriters' longer response times during the online acceptability judgment task. In contrast, L1 processing routines examined at two target points within a sentence revealed both qualitative and quantitative ERP differences between groups. Subject-verb number mismatches elicited a robust N400 effect in attriters but not in native-controls, reflecting attriters' stronger expectations for agreement between a sentence-initial NP and the verb, likely as a result of English word-order

influence. Attriters also differed from native-monolinguals in the sentence-repair processes indexed by the late posterior P600 (Hagoort and Brown, 2000; Carreiras et al., 2004; Molinaro et al., 2008). Interestingly, the late P600 was larger (i.e., more similar to native-monolinguals) in attriters with more frequent L1-Italian use.

As the experimental sentences in Kasparian, Vespignani, and Steinhauer tested combinations of both local- and non-local number agreement mismatches between three inflected constituents (noun, verb, and modifier), it seems likely that more complex morphosyntactic manipulations resulted in greater processing differences between attriters and native-monolinguals, compared to Bergmann et al. (2015). The present study aims to more directly examine L1 changes induced by the L2 grammar by testing the real-time processing of complex linguistic structures that operate differently in Italian and English, namely relative clause constructions.

Relative Clause Processing

The comprehension of relative clauses has been studied extensively across languages, with both offline and online measures. These studies have generally demonstrated that <u>subject</u> relative clauses (e.g., *The <u>reporter</u> that attacked the senator admitted the error*) are easier to process than <u>object</u> relative clauses (e.g., *The <u>reporter</u> that the senator attacked admitted the error*) in most languages (Schriefers et al., 1995; De Vincenzi, 1996; Hagoort and Brown, 2000; Friederici et al., 2001; Traxler et al., 2005; but see Carreiras et al., 2010 for an opposite preference in Basque). In the comprehension of temporarily ambiguous subject-first and object-first sentences, the initial tendency is to disambiguate the sentence toward a subject-first reading (Clifton and Frazier, 1989; De Vincenzi, 1991; Schriefers et al., 1995; Bader and Meng, 1999; Schlesewsky et al., 2000). A mismatch between the preferred/expected structure that is automatically computed online and the actual input leads to longer reading times and poorer accuracy in the less preferred condition. Several theories have been proposed to explain such processing preferences, ranging from syntactic accounts (e.g., Clifton and Frazier, 1989), working memory (WM) load (e.g., Frazier and Fodor, 1978), the simultaneous influence of syntactic and non-syntactic information (e.g., MacDonald et al., 1994), usage frequency (e.g., MacDonald et al., 1994; McRae et al., 1998), to universal complexity (e.g., MacWhinney, 1982).

A number of ERP studies have shown that unpreferred relative clause sentences create garden-path effects and require revision once the disambiguating element (e.g., number of the verb) is encountered. These processes have been associated with a centro-parietal P600 effect and/or a preceding early frontal positivity, depending on the processing difficulty involved in constructing the sentence interpretation (Mecklinger et al., 1995; Steinhauer et al., 1997; Friederici et al., 2001).

The centro-parietal P600 is an effect that has not only been elicited by outright syntactic violations (e.g., Neville et al., 1991; Hagoort et al., 1993; Friederici et al., 1996, 1999), but also by violations of structural preference in garden-path sentences (Osterhout and Holcomb, 1992, 1993; Osterhout et al., 1994), as well as in response to less expected syntactic structures

(Kaan et al., 2000). In these studies, the P600 effect has sometimes been discussed as reflecting processes of diagnosis and re-analysis or repair that are required to arrive at a well-formed sentence (see 'Diagnosis and Repair' theory by Fodor and Inoue, 1998, discussed in Friederici et al., 2001 for garden-path sentences). Larger and more prolonged P600s between 500 and 900 ms typically reflect costlier repair processes (Hagoort and Brown, 2000; Carreiras et al., 2004; Silva-Pereyra and Carreiras, 2007; Molinaro et al., 2008). The P600 response has also been associated with a mismatch between expected and actual semantic (thematic) roles assigned to NP arguments by the critical verb (Kuperberg et al., 2003; see also Hoeks et al., 2004; Kim and Osterhout, 2005). According to this view, a processing cost is incurred when semantic biases are overridden by the semantic relationships dictated by the syntactic structure of the sentence (see also Kolk et al., 2003; van Herten et al., 2005). In line with this interpretation, sentence revision and repair has been shown to be more difficult when both NPs are animate (Mak et al., 2002, 2006; Traxler et al., 2002). An ERP study of object relative clauses (Weckerly and Kutas, 1999) reported a P600 effect on both the relative clause verb and the matrix verb when thematic roles based on animacy were contradicted by the thematic roles actually assigned by the verb, that is for sentences where the inanimate (rather than the animate) noun was the subject of the verb (e.g., *The novelist that the movie inspired praised the director*...).

A somewhat earlier posterior positivity has also been discussed as a P300 component (Mecklinger et al., 1995; Friederici and Mecklinger, 1996; Steinhauer et al., 1997; Friederici et al., 2001). The P300 (specifically the P3b) has been described as reflecting a process of WM updating that may be triggered by having encountered an unexpected syntactic structure. Studies investigating garden-path effects in German object relative-clauses (Mecklinger et al., 1995; Steinhauer et al., 1997) revealed a positivity around 350 ms for participants with a high reading span – an effect that was taken to reflect a revision process that is less cognitively demanding (Friederici and Mecklinger, 1996) than revision processes which trigger a late and longer-lasting posterior P600 (see also Hagoort et al., 1999; Hagoort and Brown, 2000).

A more frontally distributed positivity (often termed "frontal P600") has also been reported for non-preferred sentence continuations or complex ambiguous sentences (Osterhout and Holcomb, 1992; Hagoort et al., 1999; Van Berkum et al., 1999; Friederici et al., 2002; Kaan and Swaab, 2003; Penolazzi et al., 2005). Similar frontal positivities have also been discussed as belonging to the P300 family (specifically a P3a; cf. Bowden et al., 2013) and reflecting surprise (Squires et al., 1975; Polich, 2007) or an attentional shift when processing an unexpected stimulus (Näätänen and Galliard, 1983).

The P600 is often accompanied by preceding negativities between 300 and 500 ms, although this pattern is more typical for morphosyntactic violations than for garden-path sentences (e.g., Friederici, 2002; Molinaro et al., 2011). In reading studies, such negativities are most often left-lateralized [i.e., the left-anterior negativity (LAN); see also Steinhauer et al., 2010 for left-temporal negativities (LTN)] and reflect

mismatches with structure-based expectations (Molinaro et al., 2011). While LANs and LTNs are typically viewed as the most likely ERP response preceding the morpho-syntactic P600, negativities may have a broader distribution near midline and are interpreted as N400 components, reflecting either additional lexical processing costs (eADM model: Bornkessel and Schlesewsky, 2006; Brouwer et al., 2012) or also mismatches with structure-based morphological expectations (Tanner et al., 2013). Most interpretations of the N400 in sentence contexts are linked to lexical processing difficulties, during either word retrieval or semantic integration (Kutas and Federmeier, 2011). Both P600 and N400 amplitudes show a gradual increase the stronger the linguistic anomaly and the more difficult the underlying processes.

Cross-Linguistic Differences in Italian and English Relative Clauses

Cross-linguistic differences in morphosyntactic properties and in semantic biases make the study of relative clause comprehension relevant for bilingual speakers, particularly when the two linguistic systems operate differently in sentence processing preferences, as is the case for English and Italian.

The two languages have been shown to differ in the cues that speakers make use of during sentence interpretation. As English has a strict word-order and a less detailed system of morphological markers, English speakers rely heavily on word-order for sentence interpretation. Conversely, Italian has a relatively free word-order and rich morphological marking system, thus number agreement and semantic information (e.g., animacy, thematic roles) are more salient cues than word-order in identifying the subject of a sentence (see "*Competition Model*"; Bates et al., 1982; MacWhinney, 1987; MacWhinney and Bates, 1989; see also Bornkessel-Schlesewsky and Schlesewsky, 2009; Bornkessel-Schlesewsky et al., 2011).

In terms of word-order, Italian relative clauses have been described as having four syntactically acceptable constructions (i.e., two different word-orders that are *both* compatible with subject- and object-first relative clauses; see **Table 1**). Sentences may follow a NP-[V-NP] structure (henceforth "V-NP") or a

---

TABLE 1 | An example of experimental sentences is provided for each condition.

Condition	
(1) V-NP-subject	Il poliziotto (S) che <u>arresta</u> i ladri (O) registra i nomi. *The policeman (S) that <u>arrests</u> the thieves (O) registers the names.*
(2) V-NP-object*	I ladri (O) che <u>arresta</u> il poliziotto (S) attendono in macchina. *The thieves (O) that <u>arrests</u> the policeman (S) wait in the car.*
(3) NP-V-subject*	Il poliziotto (S) che i ladri (O) <u>arresta</u> registra i nomi. *The policeman (S) that the thieves (O) <u>arrests</u> registers the names.*
(4) NP-V-object	I ladri (O) che il poliziotto (S) <u>arresta</u> attendono in macchina. *The thieves (O) that the policeman (S) <u>arrests</u> wait in the car.*

English translations are presented in italics. The target noun is underlined and subject/object roles are indicated in parentheses. The asterisk (*) marks those conditions which are morphosyntactic violations in English.

NP-[NP-V] structure (henceforth "NP-V") in which the relative pronoun "*che*" (= that/who) is directly followed by the second NP rather than by the verb. Although all four constructions are syntactically acceptable, NP-V-subject constructions have been described as having a low usage frequency, as they occur in poetry or songs, but less frequency in everyday Italian (Di Domenico and Di Matteo, 2009; as confirmed by acceptability ratings from native Italian speakers that we collected prior to creating our final stimuli). Given these four potential sentence constructions, the pronoun "*che*" can refer either to the subject or object of the relative clause; thus, the disambiguation of the sentence relies on semantic information and/or number agreement with the verb (Penolazzi et al., 2005; Di Domenico and Di Matteo, 2009).

In a behavioral reading study of Italian native-speakers, Di Domenico and Di Matteo (2009) tested the acceptability and processing difficulty – as reflected by reading times – of these word-orders, using reversible sentences with animate nouns, where verb number was the only disambiguating cue (e.g., *The director that criticized_sing the workers anticipated the holidays*). Results showed that the V-NP-subject construction was the most preferred[1] and was associated with the fastest reading times on the verb of the relative clause. Increased reading times registered for the V-NP-object and NP-V-subject conditions were taken as evidence of revision and integration processes, after a preferred sentence structure was initially pursued. The authors argued in favor of two processing phases: a first phase where an automatically developed sentence structure is revised, followed by a second phase further downstream where the revised interpretation and assigned thematic and syntactic roles are confirmed. Once these processes have taken place, no further reading delays were incurred on subsequent words.

In contrast, English only allows for V-NP-subject and NP-V object word-orders, whereas VP-NP-object and NP-V-subject sentences are outright syntactic violations (**Table 1**), regardless of whether the sentence interpretation is supported by semantic/thematic information and/or number agreement. It was therefore of interest to examine whether the processing routines underlying Italian relative clause comprehension may have changed as a result of prolonged daily exposure to English.

Cross-Linguistic Influence in Sentence Processing in Bilinguals

It has been widely attested that a bilingual's two languages are simultaneously active during the real-time processing of only one language. Evidence of influence of the L1 during online L2 morphosyntactic processing has been demonstrated in eye-tracking (Frenck-Mestre and Pynte, 1997) and ERP studies (e.g., Sabourin, 2003; Ojima et al., 2005; Tokowicz and MacWhinney, 2005; Kasparian et al., 2010; Foucart and Frenck-Mestre, 2011, 2012; White et al., 2012). Research has also examined the factors at play in modulating the degree of L1–L2 influence – linguistic similarity, L2 proficiency and exposure levels have been shown to affect the extent

of L1-transfer and the degree of native-like-ness in the L2 (see reviews by Kotz, 2009 and Caffarra et al., 2015). Modulations of cross-linguistic transfer, in both lexical-semantic and morphosyntactic domains, have been explained in terms of relative frequency of use and activation thresholds, with the more dominant language (generally the L1) associated with a higher baseline activation level and a better efficiency in regulating cross-linguistic competition (e.g., McDonald, 1987; MacWhinney, 1992; Kroll and Stewart, 1994; Jared and Kroll, 2001; Dijkstra and Van Heuven, 2002; Gollan et al., 2008).

In contrast, studies that have explored transfer in the *reverse* direction (L2 onto L1) have been more limited, particularly in morphosyntax (Frenck-Mestre and Pynte, 2000; Linck et al., 2009; Whitford and Titone, 2012; Timmer et al., 2014). An eye-tracking study by Dussias and Sagarra (2007) tested attachment preferences in temporarily ambiguous relative clauses (e.g., *the brother₁ of the actress₂ that? went to Boston*) in Spanish–English bilinguals with either limited or extensive L2 immersion experience, compared to native-Spanish monolingual speakers. Differences in relative clause attachment preferences were found between groups; while monolingual Spanish speakers and bilinguals with limited immersion experience reliably preferred to attach the relative clause to the first NP as Spanish speakers do (e.g., Cuetos and Mitchell, 1988; Carreiras and Clifton, 1993; Carreiras et al., 2004; Mitchell and Cuetos, 1991, Unpublished), bilinguals with extensive L2-English exposure attached the relative clause to the second NP as English speakers do (Frazier and Clifton, 1996; Carreiras and Clifton, 1999; Dussias, 2001, 2003). Interestingly, the differences between the bilingual groups held when L2 proficiency was matched. The authors take these results to support the permeability of the L1 system as a result of extensive L2 exposure. These findings can be explained within the same theoretical frameworks outlined above, only that the L2 has become the predominantly used language, rather than the L1.

It can be argued that attriters belong on the same language experience continuum as those L2 learners who have been extensively immersed in the L2, whether or not attrition consists of a more extreme shift from L1 to L2.

The Present Study

Using ERPs, the present study examined the real-time processing of four different word-orders of Italian relative clauses in a group of Italian-English adult migrants who have been predominantly exposed to English since immigration and who have unanimously reported experiencing attrition in Italian (Attriters), compared to 30 non-attriting native-speakers in Italy (Controls). In one of the earliest ERP studies of adult L1 attrition and the first to systematically manipulate a complex aspect of morphosyntax to yield a paradigm where the L1 and L2 either converge or diverge, our main aim was to determine whether there were quantitative and/or qualitative differences in L1 processing patterns in attriters, due to L2 immersion. Secondly, we studied whether L1 processing was modulated by factors such as L2-influence, L1/L2 proficiency, L1/L2 use or LoR in the L2 environment. Finally, to address the open question of whether attrition effects are more

---

[1]Acceptability means per condition were not reported in the paper but were obtained in a personal communication with the authors. The order of acceptability of the sentence conditions was #1, 4, 2, 3.

pervasive in online comprehension or in behavioral/production tasks, we compared online and offline responses.

We expected the groups to differ most on the two critical conditions [(2) V-NP-o and (3) NP-V-s], as those sentences are syntactically acceptable in Italian but not in English. If native-Controls and Attriters were to process these sentences as permissible but unpreferred (due to the mismatch with syntactic/semantic preferences and/or lower usage frequency[2]), then they may show an increased reliance on semantic cues (N400 effect for unpreferred conditions; cf. Mecklinger et al., 1995) and engage in a revision process similar to what has been documented for garden-path sentences (frontal positivity and/or P600). Instead, if Attriters were to show influence from L2-English morphosyntax, we would expect them to process V-NP-o and NP-V-s sentences as morphosyntactic violations, eliciting ERP responses that differ in latency and scalp topography from those elicited by the native-Controls. According to most authors (e.g., Molinaro et al., 2011), attriters would therefore elicit ERP responses associated with the early detection (LAN) and diagnosis/repair (robust P600) of a violation, and not show evidence of relying on semantic cues for disambiguation. Since some authors have reported that a subset of subjects elicit N400s for morpho-syntactic violations (e.g., Tanner and Van Hell, 2014), finding an N400 in Attriters would be somewhat ambiguous. In terms of individual differences, we would expect that the Attriters who differ most from Controls in their L1-processing are those individuals with higher L2 proficiency, higher L2 exposure and/or a longer LoR. Such findings would show a shift from L1-cues to L2-cues with increased L2 proficiency and exposure in adult attriters (see McDonald, 1987).

It is worth noting that our experimental design tests attriters' processing of sentences that are syntactically *correct* in their L1 (but not in their L2), rather than the typical approach of testing their responses to L1 morphosyntactic violations. Thus, while the common finding is that less exposed or less proficient speakers (usually L2 learners) elicit smaller or delayed ERP effects than native-speakers or more proficient L2 learners, in the case of the present study, we would expect the reverse, namely that L1 attriters would elicit *stronger* morphosyntactic violation effects, as a result of predominant English (L2) exposure.

MATERIALS AND METHODS

Participants

Twenty-four Italian native-speakers (14 female; *M* age: 36; Range: 25–50) who had relocated to Canada in adulthood [*M* age at immigration (AoA of English[3]): 28.2 years; Range: 18–40; *M* length of residence: 11 years; Range: 1–26] were tested at McGill University in Montreal, Canada. Attriters reported limited exposure/use of their L1-Italian (*M* daily L1 exposure: 14.92%; Range: 1–40%), and described changes or difficulties

as a result of their predominant L2-English use (*M* daily L2 exposure: 69.54%; Range: 60–96%). Thirty Italian native-speakers were tested as a control group at the University of Trento in Rovereto, Italy (17 female; *M* age: 31; Range: 25–54). They had little to no exposure to second languages (including English and Italian dialects), which we operationally defined as less than 5 h per week. All participants except one were right-handed and with no known history of neurological disorders.

Background Measures

A background questionnaire collected participants' demographic (age, gender, and education) and language information. Attriters answered additional questions about their immigration history, context and amount (in hours per week and % per day) of L1/L2 exposure and use, motivation for L1 maintenance and L2 mastery, and identity/attitudes toward each language and culture. Both groups completed four proficiency measures : (1) A written self-report measure where they rated their L1 proficiency level on a scale from 1 to 7 in listening, reading, pronunciation, fluency, vocabulary, and grammatical ability; (2) A written *C*-test (Italian version: Kraš, 2008), where they were asked to fill in the blanks in 5 short texts in which 20 words in each text had been partially deleted; (3) A written error-detection test (Kasparian, 2015), where they had to detect and correct errors in two texts; and (4) A timed verbal semantic fluency task where they were asked to produce as many vocabulary items a given semantic category within 1 min. They also completed (1) a timed reading fluency task where they silently read and answered as many true-false statements as possible in 3 min (adapted into Italian based on Woodcock et al., 2003), and (2) the letter-number-sequencing task from the Italian WAIS-IV as a measure of WM (Orsini and Pezzuti, 2013). The purpose of these tasks was to ensure that any group differences were not a result of reading speed and/or WM differences. Group means are provided in **Table 2**. Attriters scored numerically lower on all four proficiency

TABLE 2 | Group means (standard deviation) for Italian proficiency and control tasks (*p*s > 0.1).

Background measures	Controls (*n* = 30)	Attriters (*n* = 24)
Self-report of proficiency (7-point scale)	7 (0)	6.87 (0.2)
Listening comprehension	7 (0)	7 (0)
Reading comprehension	7 (0)	7 (0)
Pronunciation	7 (0)	6.96 (0.2)
Fluency	7 (0)	6.79 (0.6)
Vocabulary	7 (0)	6.63 (0.7)
Grammar	7 (0)	6.83 (0.4)
C-test (%)	96.3 (4.4)	95.2 (4.6)
Error-detection test (%)	90.0 (5.1)	89.5 (5.9)
Verbal semantic fluency (average of two categories)	23.4 (5.5)	21.5 (3.9)
Reading fluency (correct in 3 min)	71.6 (13.0)	75.3 (15.0)
Working memory		
Correct	11.2 (2.7)	11.9 (2.6)
Span	5.4 (1.1)	5.7 (1.1)

---

[2]It is not within the scope of the present study to disentangle between these views.
[3]Participants unanimously considered their AoA of English to coincide with their age of immersion into English (i.e., immigration), given that their exposure to English within the Italian school system was only minimal.

measures, though they did not differ significantly from Controls ($p > 0.1$).

Stimuli

Examples of each of the four experimental conditions are provided in **Table 1**. Each sentence began with a noun phrase (definite article + noun) which was either the subject or object of the verb in the relative clause, depending on the condition. The stimuli were based on the work of Di Domenico and Di Matteo (2009) in Italian and Mecklinger et al. (1995) in German. Noun-verb-noun triplets were created to form strong agent-patient relationships to disambiguate the sentence (e.g., attorney/convict/lawyer). Only animate nouns were used and psych verbs (fear, threaten, appreciate, love, etc.) were avoided, as they assign different theta roles (Bourguignon et al., 2012). There were no repetitions among nouns and verbs. Number was counterbalanced within each condition, such that half the sentences in each condition began with a singular subject noun, and half with a plural subject noun. Lemma frequency information for all nouns and verbs was obtained (CoLFIS database; Bertinetto et al., 2005). Both lemma frequency and length of NP1 (*M* freq.: 187.79; *M* length: 7.71) and NP2 (*M* freq.: 195.19; *M* length: 7.76) were matched across triplets ($ps > 0.1$). Sentences were nine words long; the target verb was either in fourth position (conditions 1 and 2) or sixth position (conditions 3 and 4). The final three words in the sentence were always the matrix-clause verb, a function word and a noun.

A set of 108 different sentences were constructed and realized in each of the eight conditions (four main conditions × singular/plural). Eight experimental lists were created such that, across lists, each sentence contributed equally to each condition, while no sentence was repeated within any of the experimental lists. Each participant also saw 216 filler sentences, which were part of the larger study (testing number agreement and lexical-semantic processing) and will be reported in separated papers (Kasparian and Steinhauer, 2016; Kasparian et al., 2016). Out of the total of 324 pseudorandomized stimuli (108 experimental and 216 fillers) per participant, 146 sentences (approximately 45%) were acceptable (grammatically and semantically), while 178 were expected to receive a rating of 3 or lower on a five-point rating scale (approximately 55%). Our stimuli were verified by two Italian native-speakers.

Procedure

All participants provided written informed consent prior to their participation in the study. After completing the questionnaires and behavioral tasks, participants were fitted with the EEG cap and seated in a dimly-lit, sound-attenuated booth, at approximately 80 cm from the computer monitor with a *Cedrus* seven-button RB-740 response box placed in front of them (Cedrus Corporation, San Pedro, CA, USA). Participants were instructed that their task would be to rate the acceptability of various Italian sentences on a scale from 1 (unacceptable) to 5 (perfect). We used a rating scale rather than a binary acceptability judgment task in order to better capture the range of permissibility of the relative clause constructions, which were not

outright violations in Italian. Moreover, among native-speakers, a rating scale may be more sensitive to subtle group differences than yes/no decisions. Words were presented in white 40-point Arial font characters, at the center of a black background. Each trial began with the presentation of a white fixation cross for 500 ms, followed for 200 ms by a blank screen (ISI). Each word then appeared one at a time for 300 ms (+200 ms ISI). A visual prompt ("???") followed the offset of the sentence-final word and remained on the screen until participants' button press, after which an image of the blue eye appeared at the center of the screen for a 2000 ms interval for participants to blink their eyes. The next trial began after the blinking interval. Each session lasted approximately 3 h, including setup, short breaks and cap removal. All consent forms, materials and procedures were approved by the Ethics Review Board of each institution.

EEG Recording and Analysis

The EEG was recorded continuously from 25 Ag/AgCl electrodes, 19 of which were electrodes mounted on a standard electro-cap according to the 10–20 system (Jasper, 1958), and six of which were external electrodes: four electro-oculogram (EOG) channels placed above and below the left eye (EOGV), and at the outer canthus of each eye (EOGH), as well as two reference electrodes placed on the mastoids (A1 and A2). All electrodes were referenced online to the left mastoid (A1). Impedances were kept strictly below 5 kΩ for scalp and reference electrodes, and below 10 kΩ for EOG electrodes. Signals were amplified using NeuroScan (Canada) and BrainVision (Italy) and filtered online with a band-pass filter of 0.1 to 100 Hz a sampling rate of 500 Hz. Data pre-processing and analyses were carried out using EEProbe (ANT, Enschede, Netherlands). Offline, EEG recordings were filtered with a phase-true 0.3–40 Hz band-pass filter. Trials containing artifacts due to blinks, eye-movements, and excessive muscle activity were rejected prior to averaging, using a moving-window (400 ms) standard deviation of 30 μV. On average, participants contributed 25 artifact-free trials per condition out of 27 trials, with no differences across conditions ($ps > 0.1$). One Attriter was excluded from the analysis due to exceedingly noisy trials.

Event-related potentials were analyzed separately for the V-NP and NP-V word-orders[4] and were time-locked to the onset of the verb in the relative clause. For the V-NP contrast, the baseline correction was from −200 to 0 ms. For NP-V conditions, the baseline was set at 0 to 1200 ms, due to

---

[4]Our rationale for this decision was based on several reasons. First, the word-orders differed between the first pair of conditions (V-NP) and the second pair of conditions (NP-V), resulting in differences in the sentence context that appeared prior to the relative clause verb (where we time-locked our analyses). As reported, we also used a different baseline correction for the NP-V analysis. Similarly, representative time windows to best capture the relevant ERP effects differed between the two word-orders and would have required to introduce additional time windows. Including 'word-order' (V-NP; NP-V) as a factor with relative clause type (subject, object) and group (attriters, controls) along with time-window and topographical factors would have resulted in very complex ANOVA that would most likely result in many significant interactions that would distract from the actual patterns rather than clarify them. For the sake of clarity, we believe reporting the two types of violations in separate sections is the only feasible way.
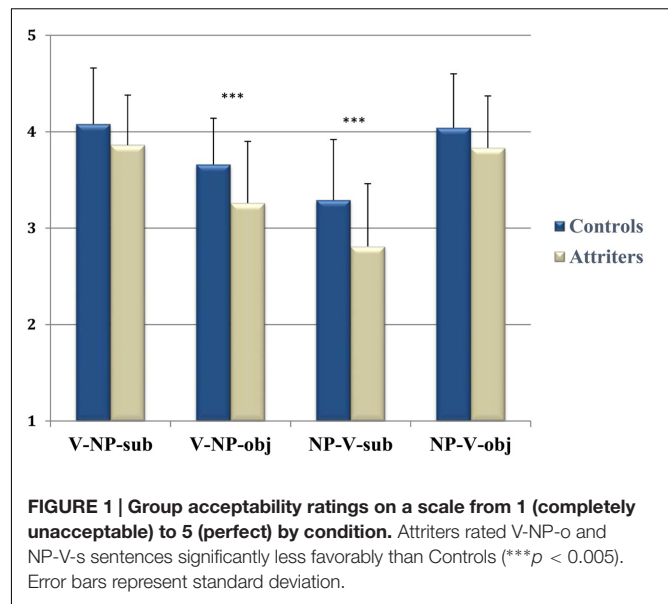
early differences in Attriters that created a baseline problem[5]. ERPs were quantified in time-windows corresponding to each component of interest, based on visual inspection of the data. For V-NP analyses, the time-windows were: (1) 300–400 (LAN/N400); (2) 650–850 (P600); (3) 850–1050 (late P600). For NP-V analyses: (1) 300–400 (LAN/N400); (2) 550–650 (frontal positivity); (3) 650–900 (P600); and (4) 900–1050 (late P600).

Repeated-measures ANOVAs were performed separately for 4 midline electrodes (Fz, Cz, Pz, and Oz) and 12 lateral electrodes (F3/4, C3/4, P3/4 and F7/8, T3/4, T5/6). Global ANOVAs for the midline sites included within-subject factors *Condition* (C: subject, object) and *Ant-Post* (AP: anterior, central, parietal, occipital). Lateral ANOVAs additionally included factors *Hemisphere* (left, right) and *Laterality* (lateral, medial). *Group* (G: Controls, Attriters) was the between-subjects factor. Greenhouse-Geisser correction was applied to analyses with more than two levels (e.g., *AP*). In these cases, the corrected *p*-values but original degrees of freedom are reported. Reported analyses are restricted to the midline only, except in cases where the lateral ANOVAs revealed additional effects. *Post hoc* analyses when following up multi-level main effects or interactions in ANOVAs were not affected by *post hoc* Bonferroni corrections; all significant *post hoc* analyses remained significant after correction for multiple comparisons (cf., Keppel and Wickens, 2004). Correlations were performed between all relevant participant factors (LoR, exposure, Italian proficiency, English proficiency) and experimental data (acceptability judgments, ERP effects quantified at a representative electrode in representative time-windows). In ANOVAs and correlational analyses, we do not report non-significant results unless motivated in specific contrasts, i.e., to emphasize the absence of an effect in one group or one condition compared to another.

RESULTS

Acceptability Judgments

Acceptability ratings (1–5) for each sentence condition are shown in **Figure 1**. Overall, the acceptability rating results were in line with the findings from Di Domenico and Di Matteo (2009) where the order of acceptability was condition # 1 < 4 < 2 < 3 in Italian native-speakers. The repeated-measures ANOVA with factor *Condition* (C: 1, 2, 3, 4) and *Group* (G: Controls, Attriters) revealed a significant C main effect [$F(3,153) = 104.184$, $p < 0.0001$] and a $C \times G$ interaction [$F(3,153) = 2.60$, $p < 0.05$]. Follow-up analyses of the C main effect showed that, across both groups, conditions 1 and 4 were significantly more accepted than conditions 2 and 3 ($p < 0.0001$ for all corresponding pairwise comparisons). Moreover, *Condition* 2 was rated as more acceptable than *Condition* 3 ($p < 0.0001$), whereas the two grammatical

FIGURE 1 | Group acceptability ratings on a scale from 1 (completely unacceptable) to 5 (perfect) by condition. Attriters rated V-NP-o and NP-V-s sentences significantly less favorably than Controls (***$p < 0.005$). Error bars represent standard deviation.

conditions 1 and 4 only differed numerically from each other ($p = 0.4$). Most importantly, the $C \times G$ interaction indicated that (compared to Controls) Attriters were more likely to reject those sentences that are ungrammatical in English. That is, Attriters (provided significantly lower ratings than Controls for V-NP-o [$F(1,52) = 10.40$, $p < 0.005$] and NP-V-s [$F(1,52) = 8.434$, $p < 0.005$] conditions, while not differing on V-NP-s and NP-V-o ($ps > 0.1$). As expected, Attriters judged the two conditions that are outright grammatical violations in English as less acceptable in Italian than native-controls, suggesting influence from their L2-English grammar. In addition, higher levels of Italian-L1 exposure were significantly correlated with more positive acceptability ratings for these unpreferred conditions (V-NP-o: $r = 0.367$, $p < 0.005$; NP-V-s: $r = 0.318$, $p < 0.01$).

In line with the interpretation that Attriters treated the two critical conditions as outright morphosyntactic violations, we found that Attriters' acceptability ratings for the two unpreferred relative-clause conditions were not found to differ statistically from ratings the same participants provided in the same experimental session for outright morphosyntactic number agreement violations ($ps > 0.1$, see Kasparian et al., 2016). Conversely, the same native-Controls provided significantly higher acceptability ratings for the RC word-orders than for number agreement violations presented in the same experimental session, indicating that they did indeed consider these RC word-orders as more grammatically acceptable than the Attriters.

To better understand group differences and variability in ratings for the two critical conditions (V-NP-o and NP-V-s), participants were clustered into high-raters and low-raters by median split across all participants' ratings[6] (**Table 3**).

---

[5]The new baseline worked against our hypotheses, as the original baseline overestimated the early negativity found in Attriters (but not in Controls). Plots with the original baseline correction are provided in Supplementary Materials.

[6]We analyzed the results separately for VN orders (2) and NV orders (2) because participants were categorized as high- or low-raters based on their acceptability judgments on the 2 unpreferred conditions (i.e., V-NP-object and NP-V-subject),

TABLE 3 | Mean (standard deviation) acceptability ratings per condition by group and rater-type.

	V-NP-s	V-NP-o	NP-V-s	NP-V-o
Low-rater Controls (*n* = 10)	3.61 (0.49)	3.27 (0.19)	2.64 (0.43)	3.64 (0.50)
High-rater Controls (*n* = 20)	4.31 (0.30)	3.86 (0.27)	3.67 (0.31)	4.28 (0.41)
Low-rater Attriters (*n* = 16)	3.71 (0.39)	2.96 (0.28)	2.51 (0.34)	3.65 (0.40)
High-rater Attriters (*n* = 7)	4.21 (0.39)	3.98 (2.58)	3.53 (0.33)	4.28 (0.26)

For V-NP word-orders, a *Condition* (subject vs. object) main effect [$F(1,53) = 51.28$, $p < 0.0001$] was qualified by a significant interaction between *Condition* × *Rater Type* × *Group* [$F(1,53) = 5.913$, $p < 0.05$]. The *Group* (Controls vs. Attriters) × *Rater-Type* (High vs. Low) interaction was significant for V-NP-o sentences [$F(1,53) = 7.49$, $p < 0.01$], as "low-rater" Attriters rated the unpreferred V-NP-o condition significantly less favorably than even "low-rater" Controls. The trend followed the same direction for NP-V-s sentences, where we found a significant interaction between *Condition* × *Rater Type* [$F(1,53) = 12.57$, $p < 0.01$]. "Low-rater" Attriters rated the unpreferred NP-V-s sentences less favorably than "low-rater" Controls, although this numerical difference did not reach statistical significance ($p = 0.7$). There were no significant differences between the two "high-rater" subgroups of Controls vs. Attriters for either condition ($ps > 0.1$). The differences between the "low rater" subgroups suggest that there is more at play than individual variability among native Italian speakers.

Reaction Times

Reaction times between the onset of the prompt and participants' button-press are shown in **Figure 2**. The repeated-measures ANOVA with factors *Condition* (1, 2, 3, 4) and *Group* (G: Controls, Attriters) revealed a main effect of *Group* [$F(1,52) = 7.547$, $p < 0.008$], reflecting Attriters' overall slower response times than Controls[7]. Contrary to previous results that unpreferred (object) and uncanonical (NP-V) sentences take longer to process (as in Di Domenico and Di Matteo, 2009), differences between conditions did not reach significance ($ps > 0.1$). This may be a result of our task (i.e., acceptability rating rather than a comprehension question) or the offline nature of this measure, as participants' responses were given at the end of the sentence rather than on the target word, as is standard practice to avoid motor artifacts in the EEG.

FIGURE 2 | Group reaction times (in seconds) by condition. Attriters were consistently slower than Controls *$p < 0.05$, **$p < 0.01$. Error bars represent standard deviation.

ERP Results for V-NP-Object vs. V-NP-Subject

Grand average ERP waveform for V-NP (Object vs. Subject) conditions time-locked to the verb of the relative clause are presented in **Figure 3** (Controls) and **Figure 4** (Attriters). In Controls, unpreferred (though syntactically acceptable) object relative clauses elicited a broadly distributed N400-like negativity (300–400 ms) and late posterior P600 (850–1050 ms). In Attriters, there is no evidence of a negativity, and the P600 effect appears to have an earlier onset, larger amplitude, and broader distribution (650–1050 ms). Group differences for relevant time intervals are illustrated with topographical maps in **Figure 5**.

N400 (300–400 ms)

The global midline ANOVA between 300 and 400 ms revealed a significant $C \times G$ interaction [$F(1,52) = 7.56$, $p < 0.01$], due to the presence of a negativity in Controls [$F(1,29) = 9.78$, $p < 0.005$] but not Attriters ($ps > 0.1$). No interactions with topographical factors pointing toward a left and/or anterior scalp distribution reached significance in the lateral ANOVA ($ps > 0.1$). The negativity in response to object-relative sentences was therefore consistent with a N400 effect.

To aid in the interpretation of the functional significance of the N400, we examined ERP patterns in relation to acceptability ratings. Our hypothesis of enhanced reliance on semantic cues in Controls is supported by the finding that Controls who provided higher acceptability ratings for V-NP-o sentences (high raters) elicited a significant N400 [$F(1,19) = 12.96$, $p < 0.005$], whereas low rater Controls elicited a weak N400 that was not statistically reliable ($ps > 0.1$)[8]. The N400 was therefore associated with higher acceptability rather than with a violation

---

and it was not always the case that the same participants were "low" raters for both unpreferred orders.

[7]Attriters were not slower across the board than Controls for all experiments conducted as part of the larger study. However, they also exhibited slower response times in a study of Italian number agreement processing (Kasparian et al., 2016) which tested similarly complex sentences.

---

[8]Despite large differences between high and low rater Controls in terms of N400 amplitudes, *F*-values, and *p*-values, the "*Condition* × *Rater*" interaction did not reach significance ($p = 0.1$). Sample size differences may have played a role. However, an important point to emphasize is that we would expect the opposite if the N400 were an indicator of a violation effect.
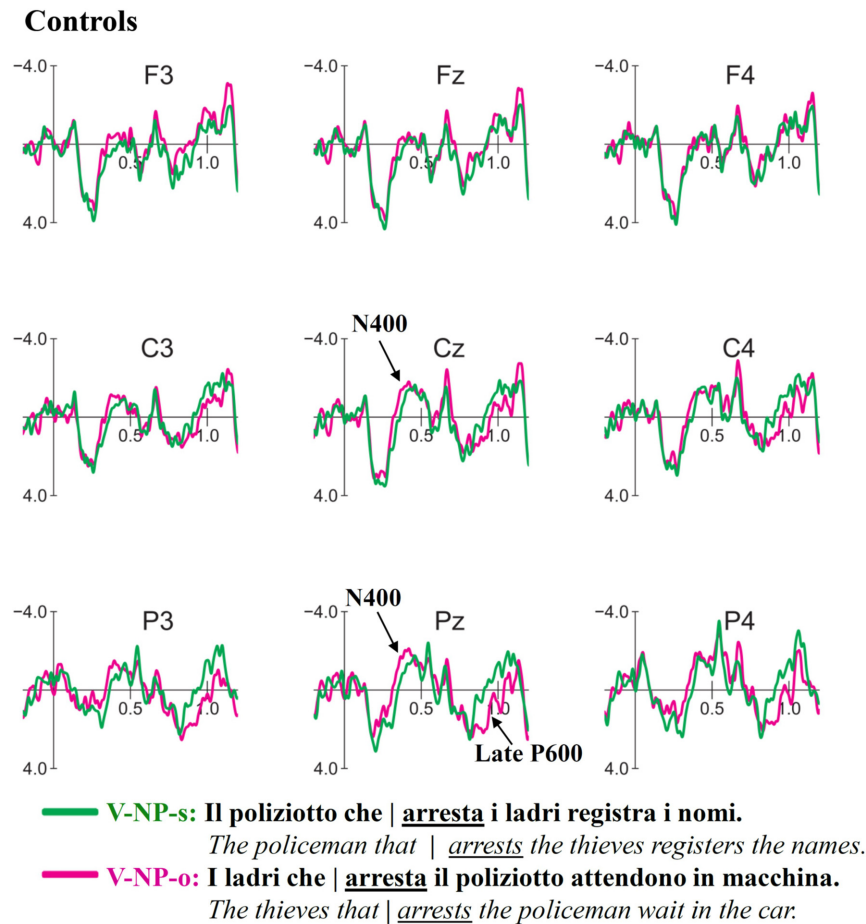
FIGURE 3 | Event-related-potentials (ERPs) elicited by the verb in response to V-NP-object sentences (pink) compared to V-NP-subject sentences (green) in Controls. Time ranges depicted on the x-axis are relative to the onset of the verb of the relative clause (0 ms). Negative values are plotted up. Controls show an N400 effect followed by a late, posterior P600 effect.

effect. In contrast, within Attriters, the N400 was absent not only in low but also high raters who did not even show a trend toward an N400 ($ps > 0.8$), suggesting an insensitivity to semantic cues in sentence interpretation[9]. These patterns are illustrated in **Figure 6**.

In line with our hypothesis that Attriters were influenced by L2-English grammar in which word-order prevails over semantic cues, correlations revealed less negative amplitudes for object-relatives at Pz in Attriters with a longer LoR ($r = 0.346, p < 0.05$) and with higher L2-English proficiency scores (C-test: $r = 0.313$, $p = 0.07$).

Early P600 (650–850 ms)

In the early time window for the P600, the midline ANOVA showed a significant $C \times AP$ [$F(3,156) = 4.56, p < 0.05$] and a marginal $C \times G$ interaction [$F(1,52) = 3.62, p = 0.06$]. Group follow-ups showed that Attriters elicited a broadly distributed P600 [$C: F(1,23) = 5.03, p < 0.05$] whereas $C \times AP$

was marginal [$F(3,69) = 3.15, p = 0.06$], but the P600 did not even approach significance in Controls ($ps > 0.6$). Attriters show an enhanced processing cost in this early P600 time-window when processing V-NP-o sentences, compared to native-Controls.

Within Attriters, a larger P600 amplitude at Pz was associated with a longer LoR ($r = 0.346 p < 0.05$) and higher L2-English proficiency scores (Semantic fluency: $r = 0.347; p < 0.05$), suggesting that increased L2 immersion and proficiency is associated with stronger morphosyntactic violation effects, as a result of L2 influence on the L1.

Late P600 (850–1050 ms)

At the midline, a significant main effect of C [$F(1,52) = 6.71$, $p < 0.01$] was qualified by a significant $C \times AP$ interaction [$F(3,156) = 8.13, p < 0.0001$], reflecting the posterior distribution of the late P600 [$F(1,52)$ at Cz < Pz < Oz]. The 3-way $C \times AP \times G$ interaction was marginal [$F(3,156) = 2.26, p = 0.08$] but reached significance in the lateral ANOVA [$F(2,104) = 6.48, p < 0.01$]. Group follow-ups revealed a significant $C \times AP$ interaction

---

[9]The difference between high rater Controls and high rater Attriters was marginal [$F(1,25) = 3.04, p = 0.09$].
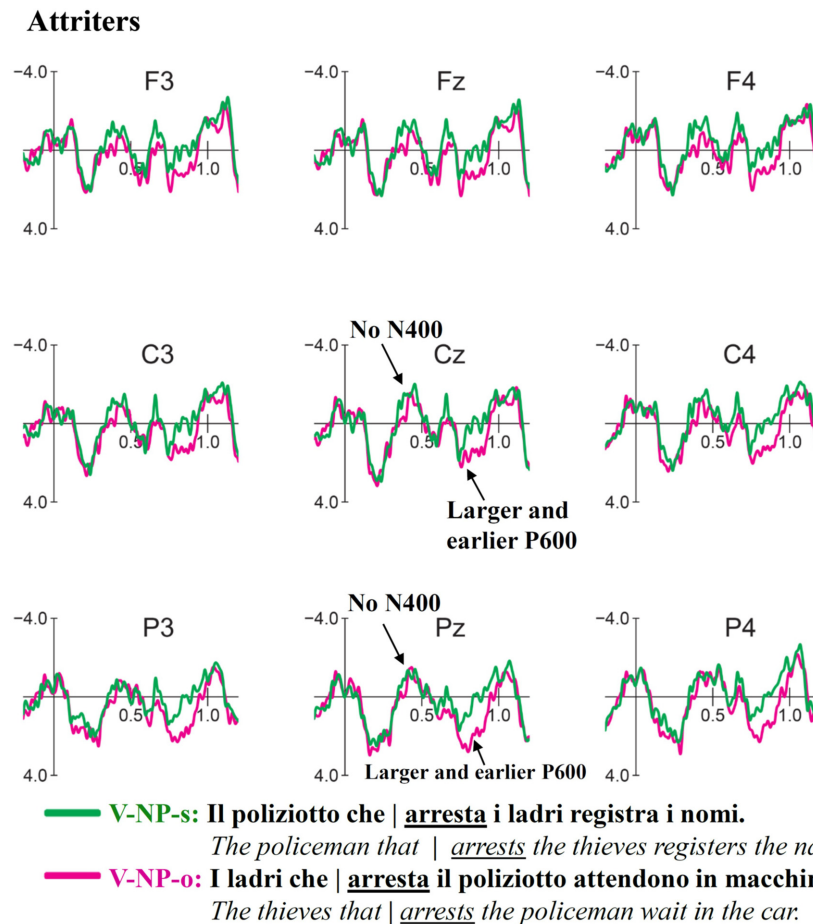
FIGURE 4 | Event-related-potentials elicited by the verb of the relative clause in V-NP-object sentences (pink) compared to V-NP-subject sentences (green) in Attriters. Unlike Controls, Attriters do not elicit an N400 effect. The P600 effect is earlier, larger and broadly distributed in Attriters.

in Controls [midline: $F(3,87) = 10.16$, $p < 0.0005$; lateral: $F(3,87) = 17.91$, $p < 0.0001$] but not Attriters ($ps > 0.1$). In Attriters, only a main effect of C was marginally significant at midline sites [$F(1,23) = 3.26$, $p = 0.08$]. Thus, overall, Attriters elicited a weaker and more broadly distributed P600 than Controls in this later time-window.

Time-Window Analysis of P600

To investigate whether the groups differed significantly in the latency of the P600, we conducted an additional analysis including factor time-window (TW) comparing the two time-windows reported above (i.e., 650–850 vs. 850–1050 ms). The midline ANOVA revealed a $TW \times C \times G$ interaction [$F(1,52) = 9.01$, $p < 0.005$], which was driven by a $TW \times C$ interaction in Controls [$F(1,29) = 13.81$, $p < 0.005$] but not Attriters ($p > 0.1$). The lateral ANOVA showed that the distribution also differed across groups and TWs [$TW \times C \times AP \times G$: $F(2,104) = 3.82$, $p < 0.05$]. Group follow-ups revealed a significant $TW \times C \times AP$ interaction in Controls [$F(2,58) = 9.45$, $p < 0.005$], whereas no interactions with factor TW reached significance in Attriters

($ps > 0.1$). This analysis further supported the finding that the P600 differed in latency and distribution between groups, with Attriters showing a more robust, earlier and more broadly distributed P600 effect for V-NP-o sentences than Controls.

ERP Results for NP-V Subject vs. NP-V Object

Grand average ERP waveforms for NP-V (Subject vs. Object) conditions time-locked to the verb of the relative clause are presented in **Figure 7**. In Controls (**Figure 7**), unpreferred (though syntactically acceptable) subject relative clauses with this word-order elicited only weak differences compared to the object relative clause: a small frontal positivity visible at Fz (550–650 ms) was followed by a small posterior P600 beginning around 700 ms. In contrast, Attriters (**Figure 8**) elicited a large negativity that extended to frontal sites (300–400 ms), a numerically larger fronto-central positivity (550–650 ms) and a larger, earlier and seemingly less posterior P600 effect than Controls. Comparing both conditions in each of the two groups indicates that the English violation
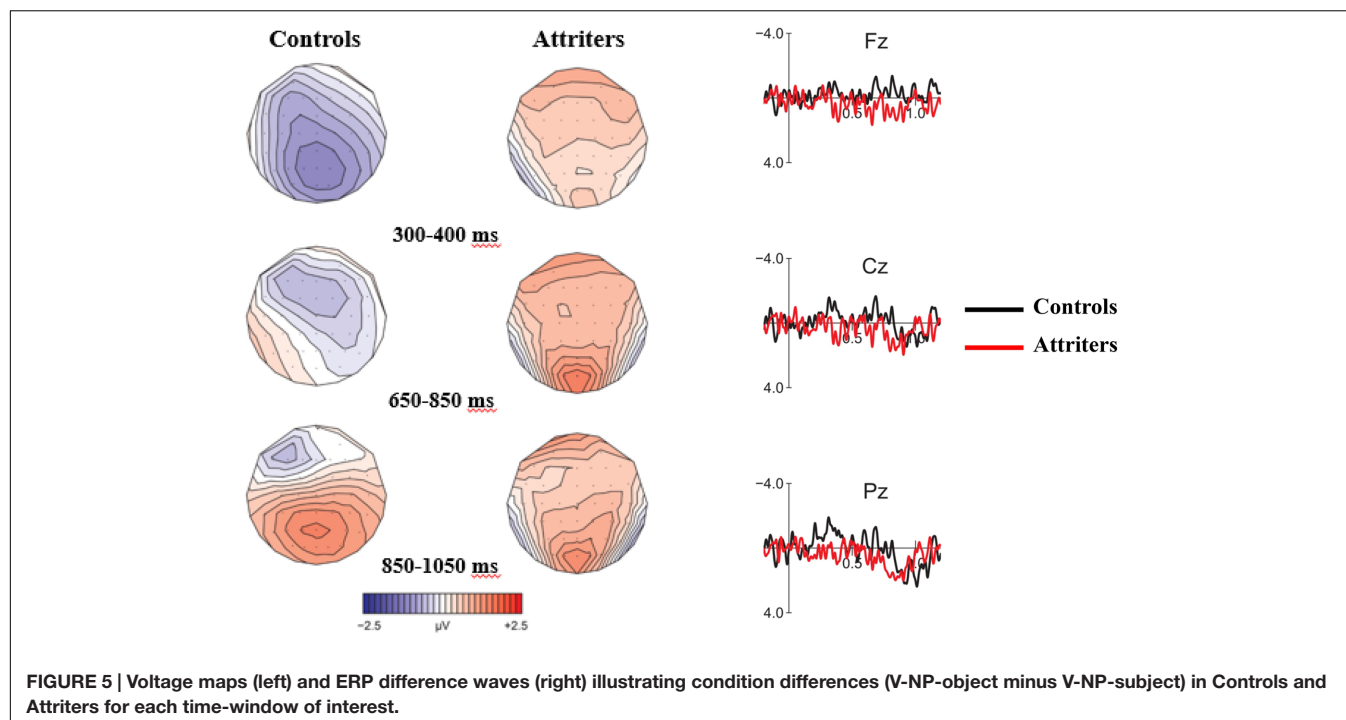
FIGURE 5 | Voltage maps (left) and ERP difference waves (right) illustrating condition differences (V-NP-object minus V-NP-subject) in Controls and Attriters for each time-window of interest.
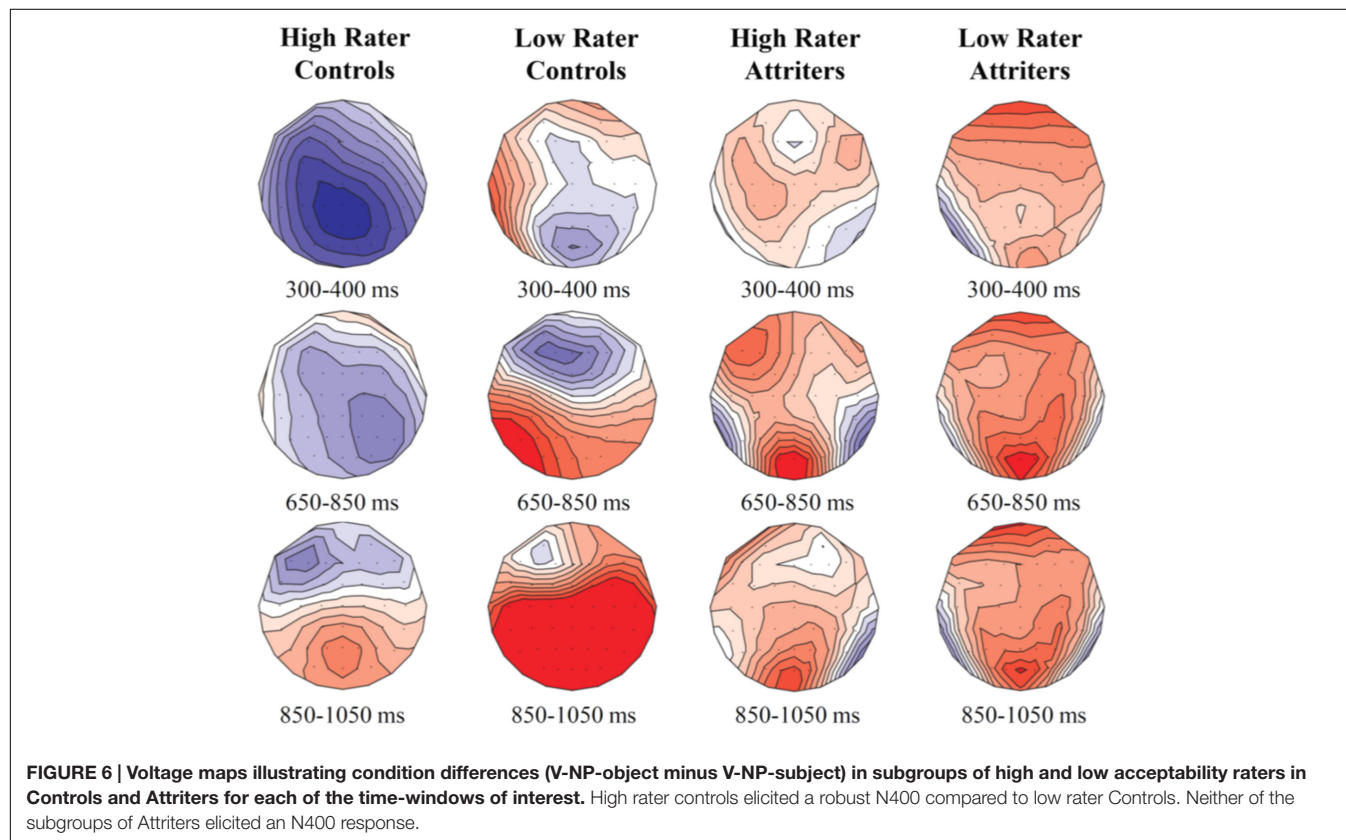

FIGURE 6 | Voltage maps illustrating condition differences (V-NP-object minus V-NP-subject) in subgroups of high and low acceptability raters in Controls and Attriters for each of the time-windows of interest. High rater controls elicited a robust N400 compared to low rater Controls. Neither of the subgroups of Attriters elicited an N400 response.

condition (NP-V-s) in Attriters is the condition that stands out when all four ERP waves are plotted together (**Figure 9**).

N400 (300–400 ms)

The midline ANOVA revealed a significant main effect of C [$F(1,51) = 6.12$, $p < 0.05$] and a marginal $C \times AP$ interaction

Controls

NP-V-s: **Il poliziotto che i ladri | arresta registra i nomi.**
The policeman that the thieves | arrests registers the names.

NP-V-o: **I ladri che il poliziotto | arresta attendono in macchina.**
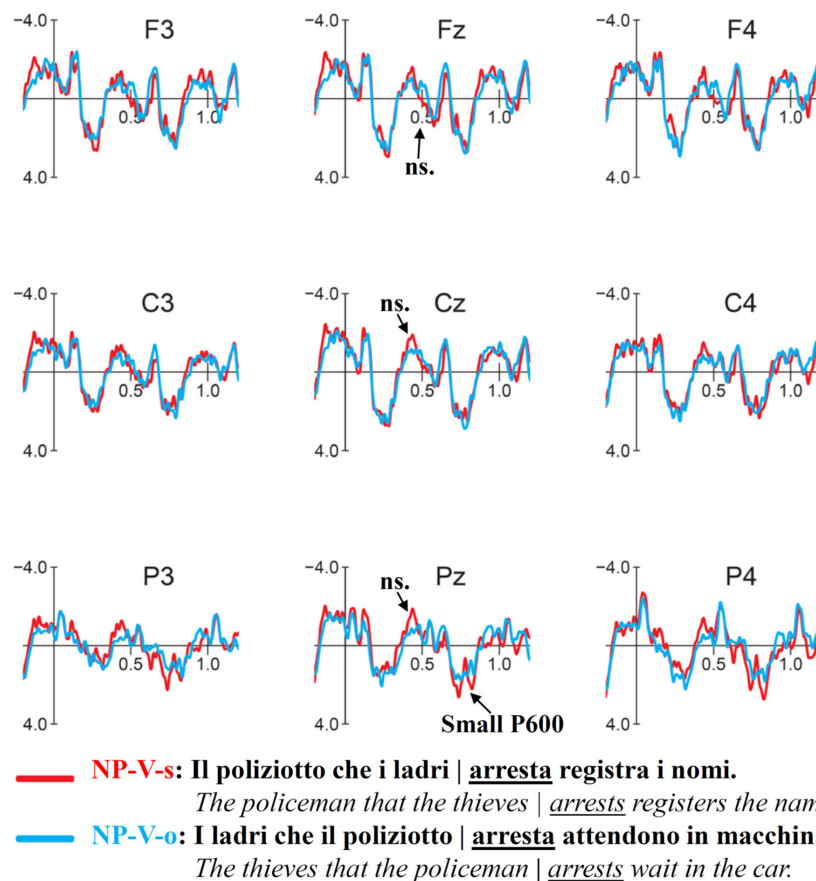*The thieves that the policeman | arrests wait in the car.*

FIGURE 7 | Event-related-potentials elicited by the verb of the relative clause in NP-V-subject sentences (red) compared to NP-V-object sentences (blue). Controls only elicited a small posterior P600 effect in response to NP-V-subject sentences.

$[F(3,153) = 2.87, p = 0.06]$. The lateral ANOVA additionally showed a marginal $C \times G$ interaction $[F(1,51) = 3.13, p = 0.08]$, which when followed-up demonstrated a negativity in Attriters $[F(1,22) = 4.31, p < 0.05]$ but not Controls $(p > 0.1)$.

Frontal Positivity (550–650 ms)

On the midline, a significant main effect of C $[F(1,51) = 13.30, p < 0.001]$ was qualified by significant $C \times AP$ $[F(3,153) = 5.86, p < 0.01]$ and $C \times G$ interactions $[F(1,51) = 4.11, p < 0.05]$. Follow-ups confirmed that the positivity in Attriters was frontal in distribution $[C \times AP: F(3,66) = 4.69, p < 0.05;$ Fz > Cz > Pz$]$ and robust $[C: F(1,22) = 11.51, p < 0.005]$ compared to Controls $(ps > 0.1)$. Rater-type (low vs. high) did not modulate the frontal positivity, and the most relevant factor was Group.

P600 (650–900 ms)

The midline ANOVA yielded significant $C[F(1,51) = 6.13, p < 0.05]$ and $C \times AP$ effects $[F(3,153) = 4.26, p < 0.01]$, reflecting the prominence of the positivity at Pz. The interaction between $C \times G$ was also significant $[F(1,51) = 4.56, p < 0.05]$, which when followed up revealed a C main effect in Attriters

$[F(1,22) = 7.95, p < 0.01]$ but not Controls $(p > 0.1)$. The interaction between $C \times AP \times G$ did not reach significance $(p > 0.1)$. Note that Controls showed no indication of a parietal P600[10].

Late P600 (900–1050 ms)

Unlike for V-NP conditions, the late P600 effect elicited in the NP-V subject condition was statistically shared by Controls and Attriters, as interactions with G did not reach significance $(ps > 0.1)$. A significant $C \times AP$ interaction $[F(3,153) = 3.79, p < 0.05]$ pointed to the posterior distribution of the positivity (Fz: $p = 0.9$; Cz: $p = 0.3$; Pz: $p < 0.05$; Oz: $p < 0.05$).

---

[10]The $C \times AP$ interaction in the global ANOVA (in absence of a significant $C \times AP \times G$ interaction) may be interpreted as an indication that even the Controls may also have some kind of a significant P600, which, however, was restricted to PZ. Alternatively, the presence of a $C \times AP$ interaction in the Controls can also be due to a relative frontal negativity, i.e., what's shared across groups is simply the gradient of 'more positive' potentials at more posterior electrodes (but in absence of a true P600). The latter pattern was what we found in a follow-up analysis at Pz. A marginal $C \times G$ interaction was found at Pz $[F(1,51) = 3.15, p = 0.08]$, further supporting a more robust P600 effect and a more laborious revision in Attriters $[F(1,22) = 11.75, p < 0.005]$ than Controls $(p > 0.1)$.
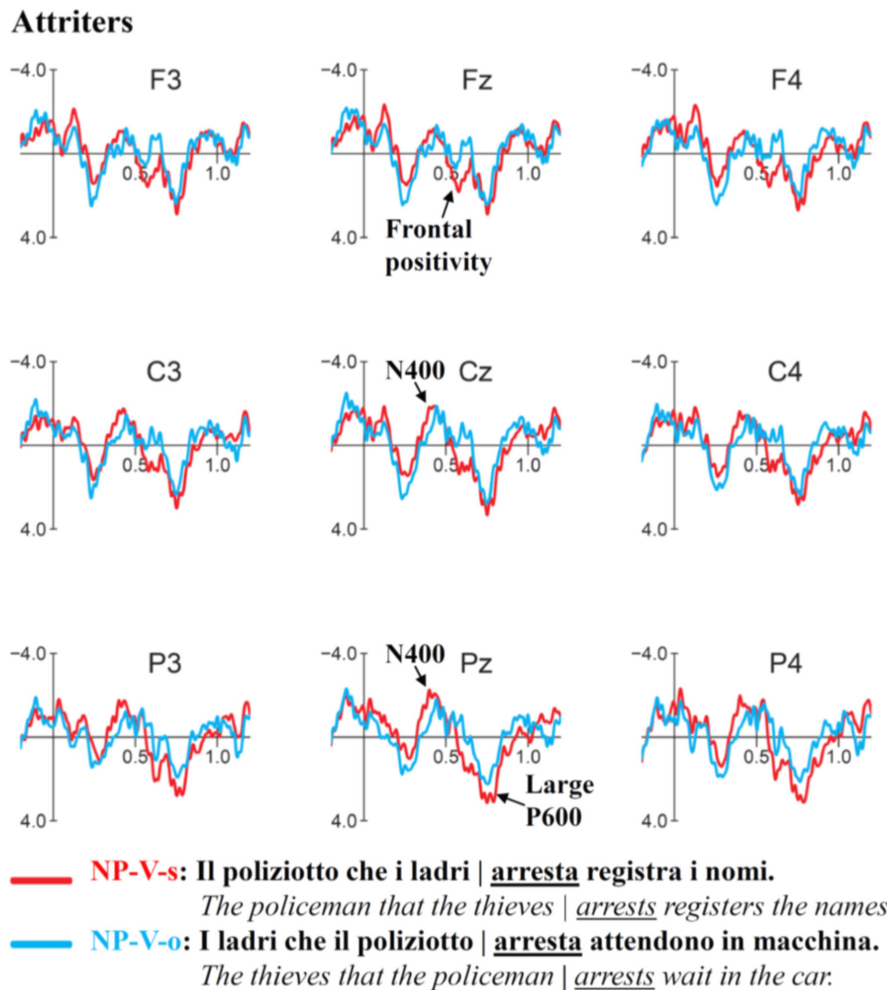
FIGURE 8 | Event-related-potentials elicited by the verb of the relative clause in NP-V-subject sentences (red) compared to NP-V-object sentences (blue). Attriters elicited an N400-like negativity, followed by an early fronto-central positivity and a large P600 effect for NP-V-subject sentences.

DISCUSSION

The present study examined the real-time L1 processing in adult Italian-migrants who had been predominantly exposed to English since immigration to Canada and who unanimously reported experiencing attrition in Italian, compared to non-attriting native-speakers still living in Italy. Our aim was to determine whether qualitative and/or quantitative differences would be found in the processing of complex relative clause constructions, due to cross-linguistic influence from the L2.

We expected Attriters to process Italian relative clause sentences whose structure would be ungrammatical in English (V-NP-object and NP-V-subject) as morphosyntactic violations, despite the presence of semantic cues to aid in the disambiguation of thematic roles. We were interested in whether such effects would also be present in Attriters' behavioral performance, and whether ERP responses would be modulated by factors such as proficiency, exposure and LoR.

Acceptability Ratings

Our main finding was that the critical conditions (V-NP-o and NP-V-s) were rated as outright morphosyntactic violations by Attriters but not by Controls. First, in native-monolinguals, the order of acceptability of the four word-order conditions [(i.e., 1, 4, 2, 3) in order of decreasing acceptability] was the same as in a previous study by Di Domenico and Di Matteo (2009), although the results of the two studies cannot be compared directly due to a difference in judgment scale and given that the stimuli in Di Domenico and Di Matteo were reversible sentences with two animate nouns, where verb number was the only disambiguating cue (e.g., *The director that criticized*$_{-sing}$ *the workers anticipated the holidays*). Given that we introduced a semantic bias in our sentences to disambiguate the agent of the verb (e.g., policeman/arrest/thief), it may be that our sentences were more readily acceptable by native-Controls.

Attriters, contrary to Controls, provided significantly lower acceptability ratings for V-NP-o and NP-V-s sentences. Crucially, the groups did not differ in their acceptability judgments of
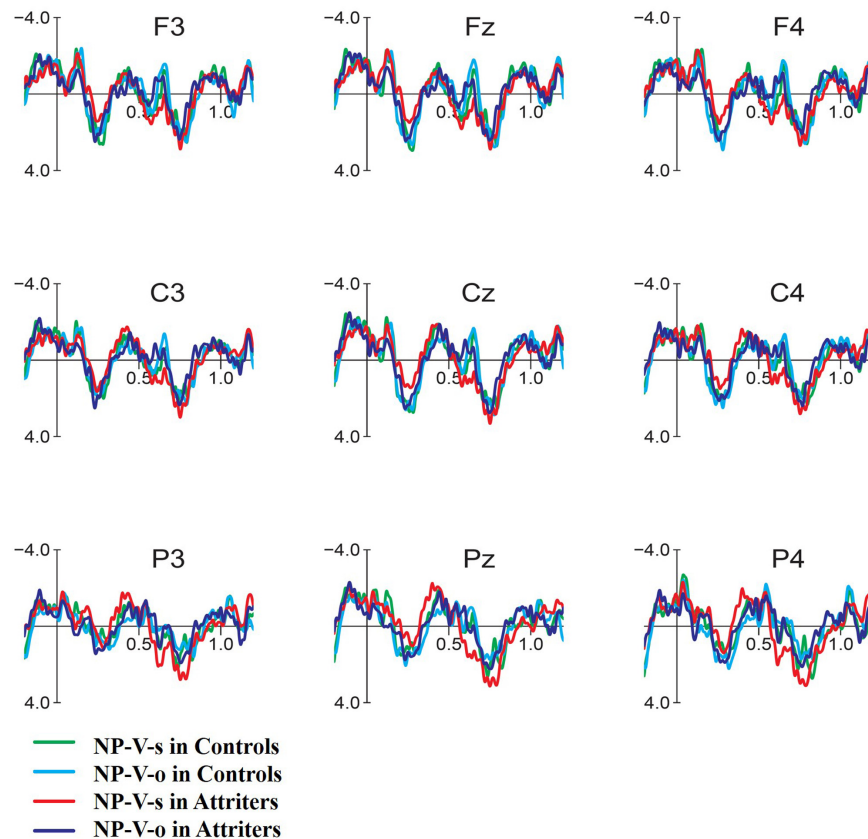
FIGURE 9 | Event-related-potentials elicited by the verb of the relative clause in each of NP-V conditions (subject and object) in Controls and Attriters. The NP-V-subject condition in Attriters (red) is the condition that differs most from other conditions.

V-NP-s and NP-V-o sentences which are syntactically acceptable in both Italian and English. Further evidence that Attriters treated V-NP-o and NP-V-s sentences as morphosyntactic violations comes from the finding that their ratings for these sentences did not differ statistically from ratings the same participants gave in response to Italian number agreement violations during the same experimental session. In contrast, Controls rated V-NP-o and NP-V-s sentences higher than the number agreement violations (see Kasparian et al., 2016). These results suggest cross-linguistic influence from English (L2) word-order during Italian (L1) sentence-reading. Given that acceptability judgments were provided at the end of each sentence and may not reflect online differences occurring at the critical sentence positions, it was of interest to determine whether and how these group differences would be reflected in real-time ERP responses.

Processing of V-NP Sentences

Our ERP findings were in line with the acceptability judgment results and demonstrate that the rating differences between groups resulted from online processing differences at disambiguating target words. In response to V-NP-object Italian sentences which are unpreferred compared to subject relative clauses but still syntactically acceptable, Controls showed an N400 effect between 300 and 400 ms at the disambiguating

verb, indicating that they were sensitive to the semantic cues that served as extra disambiguating information to identify the subject of the sentence (Penolazzi et al., 2005; Di Domenico and Di Matteo, 2009). This pattern is reminiscent of the findings of a German ERP study by Mecklinger et al. (1995), who also observed an enhanced N400 for non-preferred object relative structures, but only when semantic cues conflicted with initial parsing preferences, and only in their group of 'fast comprehenders.' German, like Italian (but unlike English), also has a free word-order, and disambiguation of relative clauses depends on verb inflection and semantic cues (see Bornkessel-Schlesewsky and Schlesewsky, 2013, for a discussion of cross-linguistic differences).

Our finding of a larger N400 with *higher* acceptability judgment ratings (numerically even within Controls) further supported the view that the elicitation of the N400 was associated with more favorable responses and therefore did not index a violation effect. The N400 is reduced in subject relative-clauses where the verb (e.g., *arrests*) is both semantically primed by its preceding noun (e.g., *policeman*) *and* represents an action compatible with this preceding noun's assumed theta role as an actor/agent. Conversely, in sentences that begin with the object (e.g., *thieves*), the enhanced N400 reflects that – despite a likely semantic priming effect – the verb may still be less expected (Van

Petten and Kutas, 1991; Federmeier and Kutas, 1999; Kuperberg, 2007), as the verb violates the thematic role that had been computed online based on the first noun.[11] In line with this, the ERP study on Italian subject/object constructions by Penolazzi et al. (2005) did not find an N400 effect preceding the reported positivities (P300 and P600), possibly due to the reversibility of their thematic roles (e.g., grandfather/kiss/child). This absence of an N400 is entirely in line with Mecklinger et al.'s (1995) findings for their 'neutral' (i.e., reversible) condition that lacked semantic biases.

Following the N400 effect, Controls also showed a late, posterior P600 response to V-NP-o relative to V-NP-s sentences. The relatively long latency of the P600 (compared to Attriters, see below) may partly be due to an ongoing N400 (i.e., the two components may have canceled each other out; cf. Steinhauer and Drury, 2012). It may also be linked to the specific type of garden-path effect involved to repair the input. As described in Penolazzi et al. (2005), even when the reader detects an unpreferred construction at the verb in the V-NP-o condition and attempts to revise it, the last constituent (subject) is not yet available; the expectation for the incoming NP to be assigned the subject role and the maintenance of an only partially constructed sentence representation in WM could incur a cognitive load and incur in a delay of syntactic integration processes, resulting in a late P600.

Attriters qualitatively and quantitative differed from Controls in their ERP responses. First, V-NP-o sentences did not elicit an N400 effect, contrary to Controls. In this respect, Attriters differed from Controls as a group overall, as even high rater Attriters did not show an N400 effect. In addition, correlations revealed less negative amplitudes in Attriters with higher L2-English proficiency scores and a longer LoR. This finding is in line with the argument that the N400 in Controls did not reflect a violation effect, as Attriters with more L2-English immersion were significantly less likely to show a negativity in the N400 time-window. These results suggest that Attriters were not sensitive to semantic cues (i.e., non-reversible agent-patient roles) to guide thematic role assignment during online sentence processing. Instead they seemed influenced by their L2-English grammar in which word-order is the most salient cue for sentence interpretation (Bates et al., 1982; McDonald, 1987; MacWhinney and Bates, 1989; Bornkessel-Schlesewsky and Schlesewsky, 2013). A similar shift in processing preferences has been shown in L2-dominant speakers in previous work by Dussias and Sagarra (2007) for relative clause attachment. We argue that L2-English immersion (with infrequent exposure to the Attriter's L1-Italian) leads to changes in expectations pursued during online language processing. That is, due to the influence of strict English word-order, Attriters likely have a stronger expectation for the first

noun of the sentence to be the subject NP (possibly to be assigned the role of Agent) such that they rely more heavily on word-order than on semantic information, compared to native-Italian monolingual speakers (Molinaro et al., 2011). As a consequence, the violation on the verb would be processed primarily as a morpho-syntactic agreement violation, whereas the semantic-thematic mismatch may be less salient than in the Control group[12].

In line with this interpretation, the second ERP difference between Attriters and Controls was that Attriters showed an earlier, stronger and more broadly distributed P600 effect for V-NP-object sentences, which we interpreted as a reflection of a stronger anomaly (and possibly higher processing costs) compared to Controls. The finding of larger P600 amplitudes in participants who gave lower acceptability ratings to V-NP-o sentences supported our interpretation of a stronger violation effect in Attriters overall.

The P600 has been interpreted in various ways, perhaps most prominently as an index of morpho-syntactic error diagnosis and structural sentence re-analysis/repair (e.g., Fodor and Inoue, 1998, discussed in Friederici et al., 2001), with larger amplitudes reflecting a larger syntactic processing difficulty (Hagoort and Brown, 2000; Carreiras et al., 2004; Silva-Pereyra and Carreiras, 2007; Molinaro et al., 2008). Since so-called 'semantic P600s' (rather than N400s) have been observed for sentences containing thematic role reversals, it has also been suggested that P600s may reflect the resolution of mismatches between two or more distinct (e.g., semantic and syntactic) processing streams (Kuperberg et al., 2003; see also Hoeks et al., 2004; Kim and Osterhout, 2005). When semantic expectations are contradicted by the syntactic structure of the sentence, a processing cost is incurred (Kolk et al., 2003; van Herten et al., 2005). Yet others have suggested that processes underlying the P600 may comprise the 'construction, revision, or updating of a mental representation of what is being communicated' at multiple levels (Brouwer et al., 2012).

The finding of group differences both in latency and scalp topography of the P600 suggests the involvement of different processing routines for V-NP-object sentences in Attriters compared to native-controls. In Controls, this mild late effect reflects re-analysis and repair processes. When semantic cues that reliably support the respective theta roles are accessed early on (N400) and revision toward an object-relative interpretation is not costly, Controls elicited a smaller late P600 and arrived at higher acceptability ratings. Conversely, a stronger processing cost in Attriters seems to be related to their reduced use of semantic information and stronger expectations of number agreement between the sentence-initial noun and the subsequent verb. Relying on a 'subject-first' processing strategy typical for English, Attriters encountered a morphosyntactic number

---

[11]Readers familiar with the recent debate on 'semantic P600s' and 'semantic illusions' may notice that our N400 findings (as well as those in Mecklinger et al., 1995) are problematic for most accounts proposed to explain the processing of role reversals (e.g., Brouwer et al., 2012, and other models discussed in their paper). We mention this debate only briefly, as a broader discussion is beyond the scope of our paper. See Bourguignon et al. (2012) and Bornkessel-Schlesewsky and Schlesewsky (2013) for more N400 findings in role reversals that are difficult to explain by semantic priming alone.

[12]To better assess the reliance on semantic cues for sentence interpretation, it would have been informative to test a fully balanced set of stimuli using both reversible (i.e., semantically neutral) and non-reversible (i.e., semantically biasing) sentences (see Mecklinger et al., 1995). However, given that the present study was one out of four experiments embedded into the same Italian testing session, it was not possible to present a full set of 16 conditions (8 main conditions counterbalanced in singular/plural number). For a fully balanced design including reversible sentences we hypothesize that a semantic facilitation effect would be found in Controls but not Attriters.

agreement violation on the verb, triggering the typical profile of a substantial P600 violation effect and leading to a low acceptability rating at the end of the sentence (suggesting they did not successfully reanalyze the structure). Consistent with this view, larger P600 amplitudes were associated with a higher degree of L2-English proficiency and a longer LoR. That increased L2 immersion and proficiency was associated with stronger morphosyntactic violation effects in response to V-NP-object sentences strongly suggests L2 influence on L1 morphosyntax and a shift in Attriters' expectations during online sentence processing.

In fact, the latency and topographical differences in the P600 between Attriters and Controls are somewhat reminiscent of the patterns observed in the same exact participants while processing subject-verb number agreement violations during the same Italian experimental session (Kasparian et al., 2016). In response to number violations, Attriters elicited a large P600 effect beginning around 650 ms that was less posterior and shorter than in Controls. Their acceptability ratings for this outright agreement violation and the 'apparent' agreement violation in our present V-NP-object condition were also comparable. This similarity further supports our view that Attriters have a stronger subject-first preference than Controls, based on the influence of English word-order. This strong preference leads Attriters to diagnose a number mismatch between the verb and its preceding noun, even if the sentence is grammatically acceptable in Italian. Importantly, outright agreement violations in the Control group did *not* result in the P600 pattern we observe for our present V-NP-object condition. Their P600 for outright agreement violations was not delayed but started at the same time as the P600 in Attriters (i.e., around 650 ms). Correspondingly, outright violations were rated as less acceptable than the V-NP-object garden-path sentences.

A final point to discuss is whether the smaller and later P600 in Controls might be directly linked to the presence of an N400 due to component overlap (see Osterhout and Mobley, 1995; Osterhout, 1997; Steinhauer and Drury, 2012; Tanner et al., 2013, 2014; Tanner and Van Hell, 2014; Tanner, 2015). This possibility is compatible with our finding that the early portion of the N400 has a typical broad distribution with a centro-parietal maximum, whereas the later negativity (after 500 ms) is more frontal (see **Figure 4**). This is exactly the pattern one would expect if the late portion of the N400 and the early portion of a clearly more posterior P600 (significant after 850 ms) superimposed one another and canceled each other out. If component overlap is indeed the main reason for the absence of an earlier P600 (present only in Attriters), this would imply that *both* the P600 and the N400 observed in the Control group were underestimated, i.e., the actual N400 must have been even larger and must have lasted longer. If so, and given that the Attriters did not show any evidence of an N400 at all, this would illustrate just how different the processing strategies between the groups were. On the other hand, if the observed pattern was not influenced by component overlap, the finding of a substantially delayed P600 (starting around 850 ms) in the monolingual native speakers of the Control group is difficult to reconcile with the expected ERP profile for a number violation in this group (see previous paragraph) and

would, again, point to distinct processing strategies compared to the Attriters.

Processing of NP-V Sentences

The first contrast we discussed above compared the processing of the generally highly preferred V-NP-subject relative clause to a relatively difficult V-NP-object garden-path sentence. We saw that the monolingual Control group processed the latter like a garden-path and used semantic cues, whereas Attriters with strong exposure to English processed it like an outright morphosyntactic violation, as would be expected for English speakers. The second comparison of NP-V structures differs from this first comparison in various respects. First, when the disambiguating element is reached (verb), both noun phrases have already been encountered. Thus, the Control group that was shown above to use semantic cues can be expected to compute even stronger expectations based on the preliminary assignment of Actor and Undergoer to the available NPs (Bornkessel-Schlesewsky and Schlesewsky, 2013).

Second, whereas for Attriters (who employ a parsing strategy influenced by English), the present contrast is somewhat similar to the previous V-NP contrast, this may not be the case for the Controls. From the Attriters' perspective, we again compare one sentence structure that is grammatical in English (V-NP-object) to a second structure that is ungrammatical in English (V-NP-subject). If Attriters are indeed generally influenced by English parsing preferences, we would expect another instance of morphosyntactic violation effects. By contrast, from the Controls' perspective, we compare two structures that can both be described as garden path sentences. Whereas NP-V-object sentences are the preferred structure to express object relative clauses in Italian, they nevertheless constitute a non-preferred structure compared to the V-NP subject relative clauses discussed in the previous section, and should encourage Italian Controls to use semantic cues to disambiguate the structure.

The NP-V-object structure requires an Object-Subject-Verb analysis (e.g., *The thieves that the policeman arrests* …) and is (similar to English) a quite frequent construction in Italian. The NP-V-subject structure requires a Subject-Object-Verb analysis (e.g., *The policeman that the thieves arrests* …), which is not very frequent in Italian but has the potential advantage that the first NP (i.e., the referent of the relative pronoun) still serves as the subject of the relative clause (similar to the most preferred V-NP-subject structure). Nevertheless, in line with Di Domenico and Di Matteo's (2009) behavioral study, these sentences received the lowest acceptability ratings in both groups (but were comparable to those for outright violations only in the Attriters).

In response to NP-V-subject garden path sentences in Italian, Controls elicited a small, late, posterior P600 starting around 900 ms. In Attriters, however, NP-V-s sentences elicited a strong, widespread N400-like negativity, followed by a larger frontal positivity and a more robust early P600 starting around 650 ms. As a whole, this pattern is (again) compatible with the assumption that the Control group processed the difference between conditions as a garden path, whereas the Attriters processed it like an outright violation. Given that Controls'

acceptability ratings displayed substantial differences between the two conditions, the rather weak ERP differences may be somewhat surprising. However, the notion that both sentences should be viewed as garden path structures may provide some explanation. As mentioned above, in both NP-[NP-V] constructions, both subject and object NPs had already been encountered before the disambiguating verb was presented, and our materials always provided reliable semantic cues as to which NP was a plausible Actor/Agent or a plausible Undergoer/Patient. According to both the Competition Model (MacWhinney and Bates, 1989) and the eADM model (Bornkessel-Schlesewsky and Schlesewsky, 2006/2008), Italian Control subjects were expected to use these semantic cues, either to predict the theta role assignment or to support the final analysis once the verb information became available. We assume that these processes were largely the same in *both* conditions, even though the less frequent NP-V-subject condition was identified as somewhat more difficult (eliciting a small P600) and resulted in a lower rating.

In Attriters, we find substantial ERP differences between the two conditions, consisting of a negativity followed by positivities. Interestingly, while the NP-V-object condition (which is grammatical in English) did not differ from the corresponding ERPs in the Control group, the NP-V-subject condition (which is ungrammatical in English) did (**Figure 9**).

The timing and distribution of the negativity suggest a mix between a LAN and an N400, with no hemispheric differences reaching statistical significance. In the literature, both LANs and N400s have been reported for number agreement violations, and both have been linked to lexical mismatch effects based on context-based morphological predictions (Molinaro et al., 2011; Tanner and Van Hell, 2014), or to lexical retrieval problems (Brouwer et al., 2012). An N400 would also be consistent with the view that computing the thematic roles requires access to lexical-semantic information (Deutsch and Bentin, 2001; Barber et al., 2004; Molinaro et al., 2008, 2011), if we assume that, at this point in the sentence, readers must determine which of the two presented nouns is the subject of the relative clause verb. In other words, even though Attriters do not seem to have used semantic cues as soon as they were provided by the NPs, once all NPs and the verb were available, they had to assign theta roles.

The presence/absence of the negativity cannot be reduced to component overlap (Osterhout and Mobley, 1995; Osterhout, 1997; Tanner et al., 2013, 2014; Tanner and Van Hell, 2014; Tanner, 2015), as the first (frontal) positivity was *larger* in the Attriters, i.e., in the group that also elicited the preceding negativity. Rather, Attriters and Controls seem to be engaging in different processing routines. We interpret the larger frontal positivity in Attriters as a P3a component that has been associated with a surprise effect and shift in attention (Squires et al., 1975; Polich, 2007). Similar frontal positivities were previously observed in temporary subject/object ambiguities in Italian *wh* constructions for target words disambiguating the more difficult object reading (Penolazzi et al., 2005). In our study, this frontal P3a was immediately followed by a large and early P600 that is indicative of a violation effect and corresponding processing costs, similar to the V-NP contrasts discussed above.

To summarize, Attriters who, due to English influence, were hypothesized to be less sensitive to semantic/thematic cues than to word-order preferences, elicited strong ERP violation effects on the verb in both V-NP-object and NP-V-subject constructions. For the same sentences, matched monolingual Italian control subjects demonstrated weaker ERP effects that are expected for native speakers processing these types of garden-path sentences. Since the two sentence conditions are ungrammatical in English, but grammatical in Italian, the group differences strongly point to distinct processing strategies. As predicted, Attriters seemed to have adopted parsing strategies from their predominantly used English L2.

Implications for First Language Attrition

The present study investigated L1 attriters' processing of a complex aspect of morphosyntax where the L1 and L2 either converge or diverge. Interestingly, our experimental design allowed us to examine neuroplasticity in Attriters' processing routines for grammatical sentences in their L1, which happened to be ungrammatical in their L2. Similar to a few previous ERP studies (e.g., Thierry and Wu, 2007; Kasparian et al., 2010), this approach focuses on the influence of a seemingly 'irrelevant' language (English) on the presented language under investigation (Italian) and thus differs from testing morphosyntactic violations in the language presented, as is traditionally done in ERP language research. Moreover, in contrast to the few other studies using this approach to test the impact of L1 on L2, here we investigated the impact of L2 on L1. Our findings provide evidence of cross-linguistic influence from the L2 due to immersion and reduced L1 use, resulting not just in quantitative but also in qualitative changes in adult Attriters' processing patterns, contrary to the findings reported in the only other published ERP research of L1 attrition investigating a different population of attriters (Bergmann et al., 2015). The present results further support and extend those reported in our Italian number agreement study (Kasparian et al., 2016), suggesting that more complex morphosyntactic manipulations result in greater processing differences between attriters and non-attriting native-speakers. In contrast to our other study, the present sentence structures were specifically selected to maximize differences in the cues that readers could rely on in Italian vs. English (i.e., semantic cues and word-order, respectively). Although behavioral studies have shown L2 to L1 transfer to occur in instances where a grammatical L2 feature is transferred to the L1 despite its ungrammaticality in L1 (e.g., Rippert and Kuiken, 2009), to our knowledge, this is the first demonstration where the opposite is true; namely, that high proficiency in a second language acquired in adulthood may render a *grammatical* sentence in one's native language *ungrammatical* when processed in real-time.

Our ERP findings are in line with reports from eye-tracking studies of immersed L2-English speakers' processing of L1 Spanish relative clauses (Dussias and Sagarra, 2007) where differences were found in relative clause attachment preferences in bilinguals with extensive L2-immersion, compared to native-Spanish monolingual speakers and bilinguals with limited L2

exposure. Although the bilinguals in Dussias and Sagarra's study were not attriters (with limited/no exposure to the L1 in the L2 environment), their eye-tracking results parallel our ERP findings of changes in L1 processing as a result of extensive L2 exposure.

Although we are only at the very beginning of a long way to better understand the neurocognitive changes involved in first language attrition, the present results are very promising. In our opinion, they cannot be explained in terms of a mere "bilingualism effect" (i.e., a by-product of having compared monolingual Controls to bilingual Attriters), as even within Attriters, ERP responses are modulated by factors such as exposure, proficiency and LoR. These factors have been shown to modulate ERP response patterns in these same Attriters on other lexical-semantic and morphosyntactic properties, both in their L1 (Kasparian and Steinhauer, 2016; Kasparian et al., 2016) as well as in their L2 (Kasparian et al., unpublished). Interestingly, in the latter study, we showed reduced L1 activation (increased inhibition) during L2 processing in Attriters with less frequent L1 exposure/use and a longer LoR. These findings fit with frameworks of relative frequency of use and activation thresholds, where the more dominant language is associated with a higher baseline activation level and a better efficiency in inhibiting cross-linguistic competition (e.g., McDonald, 1987; MacWhinney, 1992; Kroll and Stewart, 1994; Jared and Kroll, 2001; Dijkstra and Van Heuven, 2002; Gollan et al., 2008).

In attrition research, the relationship between attrition effects observed in behavior and at the brain level is still largely unexplored. On the behavioral proficiency tasks we administered, Attriters scored numerically lower than Controls but did not differ significantly on any of the measures. However, their end-of-sentence acceptability judgments largely reflected the preferences observed during real-time sentence processing, namely that the two word-orders that are ungrammatical in English were judged as unacceptable. This fits with the argument made by Steinhauer et al. (2009) that structure-specific proficiency (rather than overall proficiency) best predicts ERP response patterns. However, this was also an interesting finding, given that in our number agreement study (Kasparian et al., 2016), we found group differences in ERP responses but not in acceptability ratings (see McLaughlin et al., 2004 for a similar finding in L2 vocabulary acquisition). The nature of the sentences may explain this discrepancy; in our number agreement experiment, we manipulated the agreement between three sentence constituents (subject, verb, and modifier), giving rise to different combinations of (dis)agreement that may have resulted in a less straightforward acceptability judgment task than in our present study. In addition, our current design directly tested a morphosyntactic area where the two languages either converge or clash. It is likely that the language areas and tasks on which Attriters differ most from native-Controls are those which tap directly into the effects of L2 influence on L1.

In sum, the present study provides evidence of neurocognitive change due to language learning in adulthood. Our results revealed both quantitative and qualitative changes in L1 morphosyntactic processing patterns of Italian native-speakers who had lived in an exclusively monolingual L1 context until adulthood. Thus, even an "entrenched" L1 grammar is subject to

change after a prolonged period of L2 immersion and reduced L1 exposure/use. As the L2 takes the lead, its acquisition and use induces changes to attriters' L1 neurocognitive processes and results in differences from non-attriting native-speakers. In the present study, we have shown that a key factor in promoting these changes in attriters is the influence of the L2 on the L1, both in terms of the language pairing and related cross-linguistic differences, as well as in terms of increasing amount of L2 exposure/use and proficiency relative to the L1.

ETHICS STATEMENT

AUTHOR CONTRIBUTIONS

KK and KS contributed equally to the experimental design of the study. KK created the experimental stimuli, programmed most parts of the experiment, recruited and tested participants in Italy and in Canada (with the help of research assistants) and conducted a large part of the data analyses. KS contributed to programming and data analyses and oversaw the project. KS and KK contributed equally to data interpretation. The manuscript was written by KK with input from KS.

FUNDING

ACKNOWLEDGMENTS

of Psychology and Cognitive Sciences, Università degli studi di Trento, Italy. We also thank research assistants Linna Jin, Kristina Maiorino, Filippo Vicari, and Paolo Zandomeneghi for their invaluable help with data collection.

REFERENCES

Ammerlaan, T. (1996). *You Get a Bit Wobbly Exploring Bilingual Lexical Retrieval Processes in the Context of First Language Attrition*. Ph.D. dissertation, Nijmegen University, Nijmegen.

Bader, M., and Meng, M. (1999). Subject-object ambiguities in German embedded clauses: an across-the-board comparison. *J. Psycholinguist. Res.* 28, 121–143. doi: 10.1023/A:1023206208142

Barber, H., Salillas, E., and Carreiras, M. (2004). "Gender or genders agreement?," in *On-Line Study of Sentence Comprehension; Eye-Tracking, ERP and Beyond*, eds M. Carreiras and C. Clifton (Brighton: Psychology Press), doi: 10.1016/S0010-9452(08)70259-4

Bates, E., McNew, S., MacWhinney, B., Devescovi, A., and Smith, S. (1982). Functional constraints on sentence processing: a crosslinguistic study. *Cognition* 11, 245–299. doi: 10.1016/0010-0277(82)90017-8

Bergmann, C., Meulman, N., Stowe, L. A., Sprenger, S. A., and Schmid, M. S. (2015). Prolonged L2 immersion engenders little change in morphosyntactic processing of bilingual natives. *Neuroreport* 26, 1065–1070. doi: 10.1097/WNR.0000000000000469

Bertinetto, P. M., Burani, C., Laudanna, A., Marconi, L., Ratti, D., Rolando, C., et al. (2005). *Corpus e Lessico di Frequenza Dell'italiano Scritto (CoLFIS)*. Pisa: Scuola Normale Superiore di Pisa.

Bornkessel, I., and Schlesewsky, M. (2006). The extended argument dependency model: a neurocognitive approach to sentence comprehension across languages. *Psychol. Rev.* 113:787. doi: 10.1037/0033-295X.113.4.787

Bornkessel-Schlesewsky, I., Kretschmar, F., Tune, S., Wang, L., Genç, S., Philipp, M., et al. (2011). Think globally: cross-linguistic variation in electrophysiological activity during sentence comprehension. *Brain Lang.* 117, 133–152. doi: 10.1016/j.bandl.2010.09.010

Bornkessel-Schlesewsky, I., and Schlesewsky, M. (2006/2008). An alternative perspective on "semantic P600" effects in language comprehension. *Brain Res. Rev.* 59, 55–73. doi: 10.1016/j.brainresrev.2008.05.003

Bornkessel-Schlesewsky, I., and Schlesewsky, M. (2009). The role of prominence information in the real-time comprehension of transitive constructions: a cross-linguistic approach. *Lang. Linguist. Compass* 3, 19–58. doi: 10.1111/j.1749-818X.2008.00099.x

Bornkessel-Schlesewsky, I., and Schlesewsky, M. (2013). Reconciling time, space and function: a new dorsal–ventral stream model of sentence comprehension. *Brain Lang.* 125, 60–76. doi: 10.1016/j.bandl.2013.01.010

Bourguignon, N., Drury, J. E., Valois, D., and Steinhauer, K. (2012). Decomposing animacy reversals between agents and experiencers: an ERP study. *Brain Lang.* 122, 179–189. doi: 10.1016/j.bandl.2012.05.001

Bowden, H. W., Steinhauer, K., Sanz, C., and Ullman, M. T. (2013). Native-like brain processing of syntax can be attained by university foreign language learners. *Neuropsychologia* 51, 2492–2511. doi: 10.1016/j.neuropsychologia.2013.09.004

Brouwer, H., Fitz, H., and Hoeks, J. (2012). Getting real about semantic illusions: rethinking the functional role of the P600 in language comprehension. *Brain Res.* 1446, 127–143. doi: 10.1016/j.brainres.2012.01.055

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

Caffarra, S., Molinaro, N., Davidson, D., and Carreiras, M. (2015). Second language syntactic processing revealed through event-related potentials: an empirical review. *Neurosci. Biobehav. Rev.* 51, 31–47. doi: 10.1016/j.neubiorev.2015.01.010

Carreiras, M., and Clifton, C. (1993). Relative clause interpretation preferences in Spanish and English. *Lang. Speech* 36, 353–372. doi: 10.1177/002383099303600401

Carreiras, M., and Clifton, C. (1999). Another word on parsing relative clauses: eyetracking evidence from Spanish and English. *Mem. Cogn.* 27, 826–833. doi: 10.3758/BF03198535

Carreiras, M., Duñabeitia, J. A., Vergara, M., De La Cruz-Pavía, I., and Laka, I. (2010). Subject relative clauses are not universally easier to process: evidence from Basque. *Cognition* 115, 79–92. doi: 10.1016/j.cognition.2009.11.012

Carreiras, M., Salillas, E., and Barber, H. (2004). Event-related potentials elicited during parsing of ambiguous relative clauses in Spanish. *Cogn. Brain Res.* 20, 98–105. doi: 10.1016/j.cogbrainres.2004.01.009

Clifton, C. Jr., and Frazier, L. (1989). "Comprehending sentences with long-distance dependencies," in *Linguistic Structure in Language Processing*, eds G. Carlson and M. Tanenhaus (Dordrecht: Springer), 273–317. doi: 10.1007/978-94-009-2729-2_8

Cuetos, F., and Mitchell, D. C. (1988). Cross-linguistic differences in parsing: restrictions on the use of the Late Closure strategy in Spanish. *Cognition* 30, 73–105. doi: 10.1016/0010-0277(88)90004-2

Datta, H. (2010). *Brain Bases for First Language Lexical Attrition in Bengali-English Speakers*. Doctoral dissertation, The City University of New York, New York, NY.

de Bot, K. (1996). "Language loss," in *Contact Linguistics: An International Handbook of Contemporary Research*, Vol. 1, eds H. Goebl, P. Nelde, Z. Stary, and W. Wölk (Berlin: Walter de Gruyter), 579–585.

De Vincenzi, M. (1991). *Syntactic Parsing Strategies in Italian: The Minimal Chain Principle*, Vol. 12. Berlin: Springer Science & Business Media. doi: 10.1007/978-94-011-3184-1

De Vincenzi, M. (1996). Syntactic analysis in sentence comprehension: effects of dependency types and grammatical constraints. *J. Psycholinguist. Res.* 25, 117–133. doi: 10.1007/BF01708422

Deutsch, A., and Bentin, S. (2001). Syntactic and semantic factors in processing gender agreement in Hebrew: evidence from ERPs and eye movements. *J. Mem. Lang.* 45, 200–224. doi: 10.1006/jmla.2000.2768

Di Domenico, A., and Di Matteo, R. (2009). Processing Italian relative clauses: Working MEMORY span and word order effects on RTs. *J. Gen. Psychol.* 136, 387–406. doi: 10.1080/00221300903266671

Dijkstra, T., and Van Heuven, W. J. (2002). The architecture of the bilingual word recognition system: from identification to decision. *Bilingualism* 5, 175–197. doi: 10.1016/j.cogbrainres.2005.04.007

Dussias, P. E. (2001). "Bilingual sentence parsing," in *One Mind, Two Languages: Bilingual Language Processing*, ed. J. L. Nicol (Malden, MA: Blackwell), 159–176.

Dussias, P. E. (2003). Syntactic ambiguity resolution in L2 learners: some effects of bilinguality on L1 and L2 processing strategies. *Stud. Second Lang. Acquis.* 25, 529–557. doi: 10.1017/S0272263103000238

Dussias, P. E., and Sagarra, N. (2007). The effect of exposure on syntactic parsing in Spanish–English bilinguals. *Bilingualism* 10, 101–116. doi: 10.1017/S1366728906002847

Federmeier, K. D., and Kutas, M. (1999). A rose by any other name: long-term memory structure and sentence processing. *J. Mem. Lang.* 41, 469–495. doi: 10.1006/jmla.1999.2660

Fodor, J. D., and Inoue, A. (1998). "Attach anyway," in *Reanalysis in Sentence Processing*, eds J. D. Fodor and F. Ferreira (Dordrecht: Springer), 101–141. doi: 10.1007/978-94-015-9070-9_4

Foucart, A., and Frenck-Mestre, C. (2011). Grammatical gender processing in L2: electrophysiological evidence of the effect of L1–L2 syntactic similarity. *Bilingualism* 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: 10.3758/s13421-015-0583-6

Frazier, L., and Clifton, C. (1996). *Construal*. Cambridge, MA: MIT Press.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpsyg.2017.00389/full#supplementary-material

Frazier, L., and Fodor, J. D. (1978). The sausage machine: a new two-stage parsing model. *Cognition* 6, 291–325. doi: 10.1016/0010-0277(78)90002-1

Frenck-Mestre, C., and Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *Q. J. Exp. Psychol.* 50, 119–148. doi: 10.1080/027249897392251

Frenck-Mestre, C., and Pynte, J. (2000). "Resolving syntactic ambiguities: cross-linguistic differences?," in *Cross-Linguistic Perspectives on Language Processing*, eds M. De Vincenzi and V. Lombardo (Dordrecht: Springer), 119–148. doi: 10.1007/978-94-011-3949-6_5

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends Cogn. Sci.* 6, 78–84. doi: 10.1016/S1364-6613(00)01839-8

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.* 22:1219. doi: 10.1037/0278-7393.22.5.1219

Friederici, A. D., Hahne, A., and Saddy, D. (2002). Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair. *J. Psycholinguist. Res.* 31, 45–63. doi: 10.1023/A:1014376204525

Friederici, A. D., and Mecklinger, A. (1996). Syntactic parsing as revealed by brain responses: first-pass and second-pass parsing processes. *J. Psycholinguist. Res.* 25, 157–176. doi: 10.1007/BF01708424

Friederici, A. D., Mecklinger, A., Spencer, K. M., Steinhauer, K., and Donchin, E. (2001). Syntactic parsing preferences and their on-line revisions: a spatio temporal analysis of event-related brain potentials. *Cogn. Brain Res.* 11, 305–323. doi: 10.1016/S0926-6410(00)00065-3

Friederici, A. D., Steinhauer, K., and Frisch, S. (1999). Lexical integration: sequential effects of syntactic and semantic information. *Mem. Cogn.* 27, 438–453. doi: 10.3758/BF03211539

Gollan, T. H., Montoya, R. I., Cera, C., and Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: aging, bilingualism, and the weaker links hypothesis. *J. Mem. Lang.* 58, 787–814. doi: 10.1016/j.jml.2007.07.001

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

Hagoort, P., and Brown, C. M. (2000). ERP effects of listening to speech compared to reading: the P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia* 38, 1531–1549. doi: 10.1016/S0028-3932(00)00053-1

Hagoort, P., Brown, C. M., and Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Lang. Cognit. Process.* 8, 439–483. doi: 10.1080/01690969308407585

Hagoort, P., Brown, C. M., and Osterhout, L. (1999). "The neurocognition of syntactic processing," in *The Neurocognition of Language*, eds C. M. Brown and P. Hagoort (Oxford: Oxford University Press), 273–317. doi: 10.1093/acprof:oso/9780198507932.001.0001

Hoeks, J. C., Stowe, L. A., and Doedens, G. (2004). Seeing words in context: the interaction of lexical and sentence level information during reading. *Cogn. Brain Res.* 19, 59–73. doi: 10.1016/j.cogbrainres.2003.10.022

Hulsen, M. (2000). *Language Loss and Language Processing: Three Generations of Dutch Migrants in New Zealand*. Doctoral dissertation, Katholieke Universiteit Nijmegen, Nijmegen.

Jared, D., and Kroll, J. F. (2001). Do bilinguals activate phonological representations in one or both of their languages when naming words? *J. Mem. Lang.* 44, 2–31. doi: 10.1006/jmla.2000.2747

Jasper, H. H. (1958). The ten-twenty electrode system of the international federation. *Electroencephalogr. Clin. Neurophysiol.* 10, 371–375.

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: 10.1016/j.cogbrainres.2004.09.012

Kaan, E., and Swaab, T. Y. (2003). Repair, revision, and complexity in syntactic analysis: an electrophysiological differentiation. *J. Cogn. Neurosci.* 15, 98–110. doi: 10.1162/089892903321107855

Kasparian, K. (2015). *The Case of the Non-Native-Like First Language: Neurophysiological Evidence of First-Language Attrition*. Doctoral dissertation, McGill University, Montreal, QC.

Kasparian, K., Bourguignon, N., Drury, J. E., and Steinhauer, K. (2010). "On the influence of proficiency and L1-background in L2 processing: an ERP study of nominal morphology in French and Mandarin learners of English," In *Poster at the Presentation at the Donostia Workshop on Neurobilingualism*, Basque Center on Cognition, Brain and Language, Donostia-San Sebastián.

Kasparian, K., and Steinhauer, K. (2016). Lexical-semantic processing of confusable words in first-language attrition and late second-language acquisition: An ERP study. *Neuropsychologia* 93, 200–217. doi: 10.1016/j.neuropsychologia.2016.10.007

Kasparian, K., Vespignani, F., and Steinhauer, K. (2016). First-language attrition induces changes in online morphosyntactic processing and re-analysis: An ERP study of number agreement in complex Italian sentences. *Cogn. Sci.* doi: 10.1111/cogs.12450 [Epub ahead of print].

Keppel, G., and Wickens, T. D. (2004). *Design and Analysis: A Researcher's Handbook*, 4th Edn. Upper Saddle River, NJ: Pearson Prentice Hall.

Kim, A., and Osterhout, L. (2005). The independence of combinatory semantic processing: evidence from event-related potentials. *J. Mem. Lang.* 52, 205–225. doi: 10.1016/j.jml.2004.10.002

Kim, J. H., Montrul, S., and Yoon, J. (2010). Dominant language influence in acquisition and attrition of binding: interpretation of the Korean reflexive caki. *Bilingualism* 13, 73–84. doi: 10.1017/S136672890999037X

Kolk, H. H., Chwilla, D. J., Van Herten, M., and Oor, P. J. (2003). Structure and limited capacity in verbal working memory: a study with event-related potentials. *Brain Lang.* 85, 1–36. doi: 10.1016/S0093-934X(02)00548-5

Köpke, B. (1999). *L'attrition de la Première Langue Chez le Bilingue Tardif: Implications Pour l'étude Psycholinguistique du Bilinguisme*. Ph.D. dissertation, Université de Toulouse-Le Mirail, Toulouse. doi: 10.1515/9783111594149.115

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. S. Schmid, B. Köpke, M. Keijzer, and L. Weilemar (Amsterdam: John Benjamins), 1–43. doi: 10.1075/sibil.28

Kotz, S. A. (2009). A critical review of ERP and fMRI evidence on L2 syntactic processing. *Brain Lang.* 109, 68–74. doi: 10.1016/j.bandl.2008.06.002

Kraš, T. (2008). *L2 Acquisition of the Lexicon-Syntax Interface and Narrow Syntax by Child and Adult Croatian Learners of Italian*. Ph.D. dissertation, University of Cambridge, Cambridge. doi: 10.1075/eurosla.10.12kra

Kroll, J. F., and Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *J. Mem. Lang.* 33, 149. doi: 10.1016/j.jecp.2008.10.004

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: challenges to syntax. *Brain Res.* 1146, 23–49. doi: 10.1016/j.brainres.2006.12.063

Kuperberg, G. R., Sitnikova, T., Caplan, D., and Holcomb, P. J. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cogn. Brain Res.* 17, 117–129. doi: 10.1016/S0926-6410(03)00086-7

Kutas, M., and Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event related brain potential (ERP). *Annu. Rev. Psychol.* 62, 621. doi: 10.1146/annurev.psych.093008.131123

Linck, J. A., Kroll, J. F., and Sunderman, G. (2009). Losing access to the native language while immersed in a second language: evidence for the role of inhibition in second-language learning. *Psychol. Sci.* 20, 1507–1515. doi: 10.1111/j.1467-9280.2009.02480.x

MacDonald, M. C., Pearlmutter, N. J., and Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychol Rev.* 101, 676–703. doi: 10.1037/0033-295X.101.4.676

MacWhinney, B. (1982). "Basic syntactic processes," in *Syntax and Semantics (1). Language Acquisition*, ed. S. Kuczaj (Hillsdale, NJ: Erlbaum).

MacWhinney, B. (ed.) (1987). "The competition model," in *Mechanisms of Language Acquisition* (Hillsdale, NJ: Lawrence Erlbaum), 249–308.

MacWhinney, B. (1992). Transfer and competition in second language learning. *Adv. Psychol.* 83, 371–390. doi: 10.1016/S0166-4115(08)61506-X

MacWhinney, B., and Bates, E. (1989). *The Crosslinguistic Study of Sentence Processing*. New York, NY: Cambridge University Press. doi: 10.2307/1423201

Mak, W. M., Vonk, W., and Schriefers, H. (2002). The influence of animacy on relative clause processing. *J. Mem. Lang.* 47, 50–68. doi: 10.1006/jmla.2001.2837

Mak, W. M., Vonk, W., and Schriefers, H. (2006). Animacy in processing relative clauses: the hikers that rocks crush. *J. Mem. Lang.* 54, 466–490. doi: 10.1016/j.jml.2006.01.001

McDonald, J. L. (1987). Sentence interpretation in bilingual speakers of English and Dutch. *Appl. Psycholinguist.* 8, 379–413. doi: 10.1017/S0142716400000382

McLaughlin, J., Osterhout, L., and Kim, A. (2004). Neural correlates of second-language word learning: minimal instruction produces rapid change. *Nat. Neurosci.* 7, 703–704. doi: 10.1038/nn1264

McRae, K., Spivey-Knowlton, M. J., and Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *J. Mem. Lang.* 38, 283–312. doi: 10.1006/jmla.1997.2543

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

Molinaro, N., Barber, H. A., and Carreiras, M. (2011). Grammatical agreement processing in reading: ERP findings and future directions. *Cortex* 47, 908–930. doi: 10.1016/j.cortex.2011.02.019

Molinaro, N., Kim, A., Vespignani, F., and Job, R. (2008). Anaphoric agreement violation: an ERP analysis of its interpretation. *Cognition* 106, 963–974. doi: 10.1016/j.cognition.2007.03.006

Montrul, S. (2008). *Incomplete Acquisition in Bilinguals: Re-Examining the Age Factor.* Amsterdam: John Benjamins, doi: 10.1017/S0272263109990088

Näätänen, R., and Galliard, A. W. K. (1983). "The orienting reflex and the N2 deflection of the ERP," in *Tutorials in Event Related Potential Research: Endogenous Components*, eds A. W. K. Galliard and W. Ritter (Amsterdam: Elsevier), 119–141. doi: 10.1016/S0166-4115(08)62036-1

Neville, H. J., Nicol, J., Barss, A., Forster, K., and Garrett, M. (1991). Syntactically based sentence processing classes: evidence from event-related brain potentials. *J. Cogn. Neurosci.* 3, 155–170. doi: 10.1162/jocn.1991.3.2.151

Ojima, S., Nakata, H., and Kakigi, R. (2005). An ERP study of second language learning after childhood: effects of proficiency. *J. Cogn. Neurosci.* 17, 1212–1228. doi: 10.1162/0898929055002436

Opitz, C. (2011). *First Language Attrition and Second Language Acquisition in a Second Language Environment.* Ph.D. dissertation, Centre for Language and Communication Studies, Trinity College Dublin, Dublin. doi: 10.1017/S0272263100006240

Orsini, A., and Pezzuti, L. (2013). *WAIS-IV. Manuale.* Firenze: Giunti OS.

Osterhout, L. (1997). On the brain response to syntactic anomalies: manipulations of word position and word class reveal individual differences. *Brain Lang.* 59, 494–522. doi: 10.1006/brln.1997.1793

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., and Holcomb, P. J. (1993). Event-related potentials and syntactic anomaly: evidence of anomaly detection during the perception of continuous speech. *Lang. Cogn. Process.* 8, 413–437. doi: 10.1080/01690969308407584

Osterhout, L., Holcomb, P. J., and Swinney, D. A. (1994). Brain potentials elicited by garden-path sentences: evidence of the application of verb information during parsing. *J. Exp. Psychol* 20:786. doi: 10.1037/0278-7393.20.4.786

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

Pallier, C. (2007). "Critical periods in language acquisition and language attrition," in *Language Attrition: Theoretical Perspectives*, eds B. Kopke, M. S. Schmid, M. Keijzer, and S. Dostert (Amsterdam: John Benjamins), 155–168. doi: 10.1075/sibil.33.11pal

Paradis, M. (2003). The bilingual Loch Ness Monster raises its non-asymmetric head again- or, why bother with such cumbersome notions as validity and reliability? Comments on Evans et al (2002). *Brain Lang.* 87, 441–448. doi: 10.1016/S0093-934X(03)00136-6

Paradis, M. (2007). "L1 attrition features predicted by a neurolinguistic theory of bilingualism," in *Language Attrition: Theoretical Perspectives*, eds B. Köpke, M. S. Schmid, M. Keijzer, and S. Dosterst (Amsterdam: John Benjamins), 9–37. doi: 10.1075/sibil.33.09par

Penolazzi, B., De Vincenzi, M., Angrilli, A., and Job, R. (2005). Processing of temporary syntactic ambiguity in Italian "who"-questions: a study with event-related potentials. *Neurosci. Lett.* 377, 91–96. doi: 10.1016/j.neulet.2004.11.074

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148. doi: 10.1016/j.clinph.2007.04.019

Rippert, A., and Kuiken, F. (2009). L2-induced changes in the L1 of Germans living in the Netherlands. *Bilingualism* 13, 41–48. doi: 10.1017/S1366728909990320

Sabourin, L., and Stowe, L. A. (2008). Second language processing: when are first and second languages processed similarly? *Second Language Research* 24, 397–430.

Sabourin, L. L. (2003). *Grammatical Gender and Second Language Processing: An ERP Study.* Doctoral dissertation, Rijksuniversiteit Groningen, Groningen.

Schlesewsky, M., Fanselow, G., Kliegl, R., and Krems, J. (2000). "The subject preference in the processing of locally ambiguous wh-questions in German," in *German Sentence Processing*, eds B. Hemforth and L. Konieczny (Dordrecht: Springer), 65–93.

Schmid, M. S. (2010). Languages at play: the relevance of L1 attrition to the study of bilingualism. *Bilingualism* 13, 1. doi: 10.1017/S1366728909990368

Schmid, M. S. (2011). *Language Attrition.* Cambridge: Cambridge University Press, doi: 10.1017/CBO9780511852046

Schmid, M. S. (2013). "First language attrition as a window to constraints on bilingual development," in *Proceedings of the Keynote Lecture at the International Symposium on Bilingualism (ISB9)*, Singapore,

Schmid, M. S., and Dusseldorp, E. (2010). Quantitative analyses in a multivariate study of language attrition: the impact of extralinguistic factors. *Second Lang. Res.* 26, 125–160. doi: 10.1177/0267658309337641

Schmid, M. S., and Köpke, B. (2011). L'attrition de la première langue en tant que phénomène psycholinguistique. *Lang. Interact. Acquis.* 2, 197–220. doi: 10.1075/lia.2.2.02kop

Schriefers, H., Friederici, A. D., and Kuhn, K. (1995). The processing of locally ambiguous relative clauses in German. *J. Mem. Lang.* 34, 499. doi: 10.1371/journal.pone.0146194

Silva-Pereyra, J. F., and Carreiras, M. (2007). An ERP study of agreement features in Spanish. *Brain Res.* 1185, 201–211. doi: 10.1016/j.brainres.2007.09.029

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. *Linguist. Approaches Biling.* 1, 1–33. doi: 10.1016/0013-4694(75)90263-1

Squires, N., Squires, K., and Hillyard, S. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalogr. Clin. Neurophysiol.* 38, 387–401. doi: 10.1016/0013-4694(75)90263-1

Steinhauer, K., and Drury, J. E. (2012). On the early left-anterior negativity (ELAN) in syntax studies. *Brain Lang.* 120, 135–162. doi: 10.1016/j.bandl.2011.07.001

Steinhauer, K., Drury, J. E., Portner, P., Walenski, M., and Ullman, M. T. (2010). Syntax, concepts, and logic in the temporal dynamics of language comprehension: evidence from event-related potentials. *Neuropsychologia* 48, 1525–1542. doi: 10.1016/j.neuropsychologia.2010.01.013

Steinhauer, K., Mecklinger, A., Friederici, A. D., and Meyer, M. (1997). Probability and strategy: an event-related potential study of processing syntactic anomalies. *Z. Exp. Psychol.* 2, 305–331. doi: 10.3758/BF03197249

Steinhauer, K., White, E. J., and Drury, J. E. (2009). Temporal dynamics of late second language acquisition: evidence from event-related brain potentials. *Second Lang. Res.* 25, 13–41. doi: 10.1177/0267658308098995

Tanner, D. (2015). On the left anterior negativity (LAN) in electrophysiological studies of morphosyntactic 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. *Bilingualism* 17, 277–293. doi: 10.1017/S1366728913000370

Tanner, D., McLaughlin, J., Herschensohn, J., and Osterhout, L. (2013). Individual differences reveal stages of L2 grammatical acquisition: ERP evidence. *Bilingualism* 16, 367–382. doi: 10.1017/S1366728912000302

Tanner, D., and Van Hell, J. G. (2014). ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia* 56, 289–301. doi: 10.1016/j.neuropsychologia.2014.02.002

Thierry, G., and Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proc. Natl. Acad. Sci. U.S.A.* 104, 12530–12535. doi: 10.1073/pnas.0609927104

Timmer, K., Ganushchak, L. Y., Ceusters, I., and Schiller, N. O. (2014). Second language phonology influences first language word naming. *Brain Lang.* 133, 14–25. doi: 10.1016/j.bandl.2014.03.004

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

Traxler, M. J., Morris, R. K., and Seely, R. E. (2002). Processing subject and object relative clauses: evidence from eye movements. *J. Mem. Lang.* 47, 69–90. doi: 10.1006/jmla.2001.2836

Traxler, M. J., Williams, R. S., Blozis, S. A., and Morris, R. K. (2005). Working memory, animacy, and verb class in the processing of relative clauses. *J. Mem. Lang.* 53, 204–224. doi: 10.1016/j.jml.2005.02.010

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: 10.1177/1367006904008003 0601

Van Berkum, J. J., Hagoort, P., and Brown, C. M. (1999). Semantic integration in sentences and discourse: evidence from the N400. *J. Cogn. Neurosci.* 11, 657–671. doi: 10.1162/089892999563724

van Herten, M., Kolk, H. H., and Chwilla, D. J. (2005). An ERP study of P600 effects elicited by semantic anomalies. *Cogn. Brain Res.* 22, 241–255. doi: 10.1016/j. cogbrainres.2004.09.002

Van Petten, C., and Kutas, M. (1991). Influences of semantic and syntactic context on open-and closed-class words. *Mem. Cogn.* 19, 95–112. doi: 10.3758/ BF03198500

Weckerly, J., and Kutas, M. (1999). An electrophysiological analysis of animacy effects in the processing of object relative sentences. *Psychophysiology* 36, 559–570. doi: 10.1111/1469-8986.3650559

White, E. J., Genesee, F., and Steinhauer, K. (2012). Brain responses before and after intensive second language learning: proficiency based changes and first language background effects in adult learners. *PLoS ONE* 7:52318. doi: 10.1371/ journal.pone.0052318

Whitford, V., and Titone, D. (2012). Second-language experience modulates first- and second-language word frequency effects: evidence from eye movement measures of natural paragraph reading. *Psychon. Bull. Rev.* 19, 73–80. doi: 10.3758/s13423-011-0179-5

Woodcock, R. W., McGrew, K. S., Mather, N., and Schrank, F. A. (2003). *Woodcock-Johnson III Diagnostic Supplement to the Tests of Cognitive Abilities.* Itasca, IL: Riverside. doi: 10.1177/003435520104400407

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: info@frontiersin.org | +41 21 510 17 00

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership