

Experimental approaches to the acquisition of information structure

Edited by

Tania L. Leal, Bradley Hoot, Stavros Skopeteas,
Joseph V. Casillas and Oksana Laleko

Published in

Frontiers in Psychology
Frontiers in Language Sciences



FRONTIERS EBOOK COPYRIGHT STATEMENT

The copyright in the text of individual articles in this ebook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this ebook is the property of Frontiers.

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or ebook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-8325-6546-9
DOI 10.3389/978-2-8325-6546-9

Generative AI statement

Any alternative text (Alt text) provided alongside figures in the articles in this ebook has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

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: frontiersin.org/about/contact

Experimental approaches to the acquisition of information structure

Topic editors

Tania L. Leal — University of Arizona, United States

Bradley Hoot — DePaul University, United States

Stavros Skopeteas — University of Göttingen, Germany

Joseph V. Casillas — Rutgers, The State University of New Jersey, United States

Oksana Laleko — State University of New York at New Paltz, United States

Citation

Leal, T. L., Hoot, B., Skopeteas, S., Casillas, J. V., Laleko, O., eds. (2025). *Experimental approaches to the acquisition of information structure*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-6546-9

Table of contents

04	Editorial: Experimental approaches to the acquisition of information structure Tania L. Leal, Bradley Hoot, Stavros Skopeteas, Joseph V. Casillas and Oksana Laleko
08	What corpus data reveal about the Position of Antecedent Strategy: anaphora resolution in Spanish monolinguals and L1 English-L2 Spanish bilinguals Cristóbal Lozano and Teresa Quesada
29	On the incompatibility of object fronting and progressive aspect in Yucatec Maya Melanie Uth, Rodrigo Gutiérrez-Bravo and Jan Fliessbach
39	L2 Arabic learners' processing of Arabic garden-path sentences: a consistent reading pattern Abdullah M. Seraye Alseraye
52	Corrigendum: L2 Arabic learners' processing of Arabic garden-path sentences: a consistent reading pattern Abdullah M. Seraye Alseraye
53	Word order and context in sentence processing: evidence from L1 and L2 Russian Natalia Slioussar and Maria Harchevnik
67	Paradigmatic and syntagmatic effects of information status on prosodic prominence – evidence from an interactive web-based production experiment in German Janne Lorenzen, Simon Roessig and Stefan Baumann
85	The development of prosodic focus marking in French Emilie Destruel, Louise Lalande and Aoju Chen
96	Prosodic focus marking in Seoul Korean-speaking children: the use of prosodic phrasing Anqi Yang, Taehong Cho, Sahyang Kim and Aoju Chen
109	L1 grammatical attrition through the acquisition of competing L2 discourse features Liz Smeets
128	Acquisition of non-contrastive focus in Russian by adult English-dominant bilinguals Tatiana Luchkina, Tania Ionin and Maria Goldshtein
148	Word order and information structure in Romeyka Nicolao Neocleous and Ioanna Sitaridou



OPEN ACCESS

EDITED AND REVIEWED BY
Xiaolin Zhou,
Peking University, China

*CORRESPONDENCE

Tania L. Leal
✉ tanialeal@arizona.edu
Bradley Hoot
✉ bhoot@depaul.edu

RECEIVED 07 April 2025

ACCEPTED 19 May 2025

PUBLISHED 17 June 2025

CITATION

Leal TL, Hoot B, Skopeteas S, Casillas JV and
Laleko O (2025) Editorial: Experimental
approaches to the acquisition of information
structure. *Front. Psychol.* 16:1607534.
doi: 10.3389/fpsyg.2025.1607534

COPYRIGHT

© 2025 Leal, Hoot, Skopeteas, Casillas and
Laleko. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Editorial: Experimental approaches to the acquisition of information structure

Tania L. Leal^{1*}, Bradley Hoot^{2*}, Stavros Skopeteas³,
Joseph V. Casillas⁴ and Oksana Laleko⁵

¹Department of Spanish & Portuguese, University of Arizona, Tucson, AZ, United States, ²Department of Modern Languages, DePaul University, Chicago, IL, United States, ³Institute for Linguistics, Georg-August University of Göttingen, Göttingen, Germany, ⁴Department of Spanish and Portuguese, Rutgers, The State University of New Jersey, New Brunswick, NJ, United States, ⁵Linguistics Program, State University of New York at New Paltz, New Paltz, NY, United States

KEYWORDS

information structure, topic/focus preposing, experimental linguistics, language acquisition, focus

Editorial on the Research Topic

Experimental approaches to the acquisition of information structure

Utterances vary in terms of their (in)felicity depending on how constituents relate to the discourse context, speakers' communicative needs, and speakers' assessment of hearers' beliefs. The study of *information structure* explores how speakers package their utterances into blocks with varying informational values, formalizing these units with notions like “focus,” “background,” or “topic” (Krifka, 2008). These categories have received substantial interest in linguistics, generating both theoretical models and experimental studies bearing on how information structure is represented and interpreted in the minds of speakers (for overviews, see Féry and Ishihara, 2016; Krifka and Musan, 2012).

Research on information structure has boomed in recent years. Our survey of related terms on Scopus spanning 1960–2024 yielded 1,879 peer-reviewed journal articles, with the bulk of scholarship published in the past decade (Figures 1a, b)¹.

Despite the surge, investigating how information structure is *acquired* remains in its early stages, with uneven coverage across populations and languages. As shown in Figure 1c, the representation of individual languages in our survey follows a power-law distribution, with 52% of the articles focusing on just four languages. Only 25% of articles in our sample include at least one term related to acquisition or bilingualism (see text footnote 1).

The contributions to this Research Topic address these lacunae by expanding the cross-linguistic scope, incorporating data from child L1 acquirers, L2 and heritage bilinguals, and contexts of societal multilingualism, and utilizing both traditional and innovative methods.

Lozano and Quesada use CEDEL2 corpus texts to examine anaphora resolution in Spanish native speakers and English-speaking Spanish learners. Their findings challenge the Position of Antecedent Strategy (Carminati, 2002) as the default strategy, showing anaphora resolution is more complex than experimental data suggests, with overt pronouns rarely used and often substituted by repeated noun phrases.

Uth et al. demonstrate, using an oral production task and a corpus study, that focus in Yucatec Maya is incompatible with progressive aspect marking. Appealing to a semantic account, they argue that progressive aspect blocks focus fronting because the marker itself functions as a type of focalization.

¹ For the list of terms, see https://github.com/jvcasillas/acquisition_information_structure.

Seraye Alseraye examines how incomplete speech representations affect processing of garden path sentences in L2 Arabic, finding faster reading times in unambiguous contexts and when disambiguating segmental information is present. Overall comprehension remained unaffected, even in the presence of incorrect disambiguating information. The study supports the “good-enough” model of language processing (Ferreira et al., 2009) among L2 learners of an understudied language.

Slioussar and Harchevnik explore how L1 Russian speakers and Mandarin Chinese L2 Russian learners process SVO and OVS word orders. Using online (reading times) and offline (sentence rating) tasks, they show that both groups benefit from given-before-new structures, although L2 learners struggle more with processing non-canonical word orders and are less sensitive to discourse constraints.

Lorenzen et al. employ a novel paradigm—an interactive reading task—to increase the ecological validity of spoken data. They examine how information status affects prosodic prominence in German, finding that paradigmatic effects appear mainly in F0, while syntagmatic effects vary across speakers and depend on the specific acoustic parameter.

Destrueel et al. investigate the acquisition of French prosody using a virtual robot-mediated picture-matching task. Unlike younger children, 7- to 8-year-olds and adults use prosody to distinguish focus from non-focus. Furthermore, this study finds subject-object asymmetries, attributed to the dominant use of syntactic strategies for subject focus in French.

Yang et al. examine how young children acquire prosodic phrasing to mark focus in Korean. Using a picture-matching task, they find that children (ages 4–5) pattern like adults in distinguishing narrow from broad focus and prefocal material, but not from postfocal material or contrastive focus. By age 11, patterns are adult-like, with acquisition speed linked to form-meaning transparency.

Smeets uses two tasks to test clitic-doubled left dislocation in Romanian, which has received less attention than other Romance languages. The finding that L1 Romanian speakers who learned L2 Italian show attrition—unlike those who learned L2 English—highlights the role of L1-L2 similarity in reshaping L1 information structure via feature reassembly.

Luchkina et al. used two aural identification tasks (with and without contexts) to investigate how English-Russian heritage bilinguals process Russian non-contrastive focus, examining constituent order and prosodic cues. While higher-proficiency heritage speakers patterned with native speakers, the group overall tended to assign focus to nouns with nuclear stress in SVO orders—unlike native speakers—which highlights the challenges external interface structures pose (Sorace, 2011).

Neocleous and Sitaridou examine information-structural reflexes of contact between VO and OV languages. Romeyka, an Asia-Minor Greek variety (VO), has coexisted alongside Turkish (OV) for centuries. As a result, left peripheral configurations like focus movement occur in a wider range of contexts than in other Greek varieties.

Each article fills the literature gaps we identified, offering directions for future research to build on. At the methodological level, a key desideratum in information structure research is

to improve the ecological validity of data, minimizing lab speech artifacts. Several contributions address this by proposing novel experimental designs (e.g., Lorenzen et al.) or combining experimental and observational research (e.g., Uth et al.). We envision future studies in which these avenues will be further pursued.

Another major challenge in studying information structure is disentangling the roles of different linguistic layers involved in its expression. The interplay between syntax and prosody in particular is central to several contributions. Destrueel et al. examine the syntax-prosody complementarity in French focus expression, while Luchkina et al. investigate how prosodic and syntactic cues contribute to focus processing in Russian. We see a continued need for such nuanced, multi-layered approaches to the cross-linguistic inventory encoding these distinctions.

Finally, studying different populations beyond literate adult monolinguals—such as naturalistic and instructed bilinguals, L1 acquirers at different stages, and speakers of vernacular varieties—is imperative to understanding how grammars vary within and across languages. Some contributions show effects on attrition (Smeets) or adaptation under language contact (Neocleous and Sitaridou), while others reveal particular processing challenges in L2 learners and other bilinguals (Slioussar and Harchevnik). These findings enrich broader discussions on how dynamic processes like acquisition and language contact shape the representation and processing of information structure across diverse linguistic systems.

Author contributions

TL: Writing – review & editing, Writing – original draft. BH: Writing – original draft, Writing – review & editing. SS: Writing – review & editing, Writing – original draft. JC: Writing – original draft, Visualization, Writing – review & editing. OL: Writing – original draft, Writing – review & editing.

Generative AI Statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Carminati, M. N. (2002). *The Processing of Italian Subject Pronouns*. Boston, MA: University of Massachusetts.
- Ferreira, F., Engelhardt, P. E., and Jones, M. W. (2009). “Good enough language processing: a satisfying approach” in *Proceedings of the 31st Annual Conference of the Cognitive Science Society*, eds. N. Taatgen, H. Rijn, J. Nerbonne, and L. Schomaker (Austin, TX: Cognitive Science Society), 413–418.
- Féry, C., and Ishihara, S. (Eds.). (2016). *The Oxford Handbook of Information Structure*. Oxford, Oxford University Press.
- Krifka, M. (2008). Basic notions of information structure. *Acta Linguist. Hung.* 55, 243–276. doi: 10.1556/aling.55.2008.3-4.2
- Krifka, M., and Musan, R. (Eds.). (2012). *The Expression of Information Structure*. Berlin: De Gruyter Mouton.
- Sorace, A. (2011). Pinning down the concept of “interface” in bilingualism. *Linguist. Approach. Bilingual.* 1, 1–33. doi: 10.1075/lab.1.1.01sor

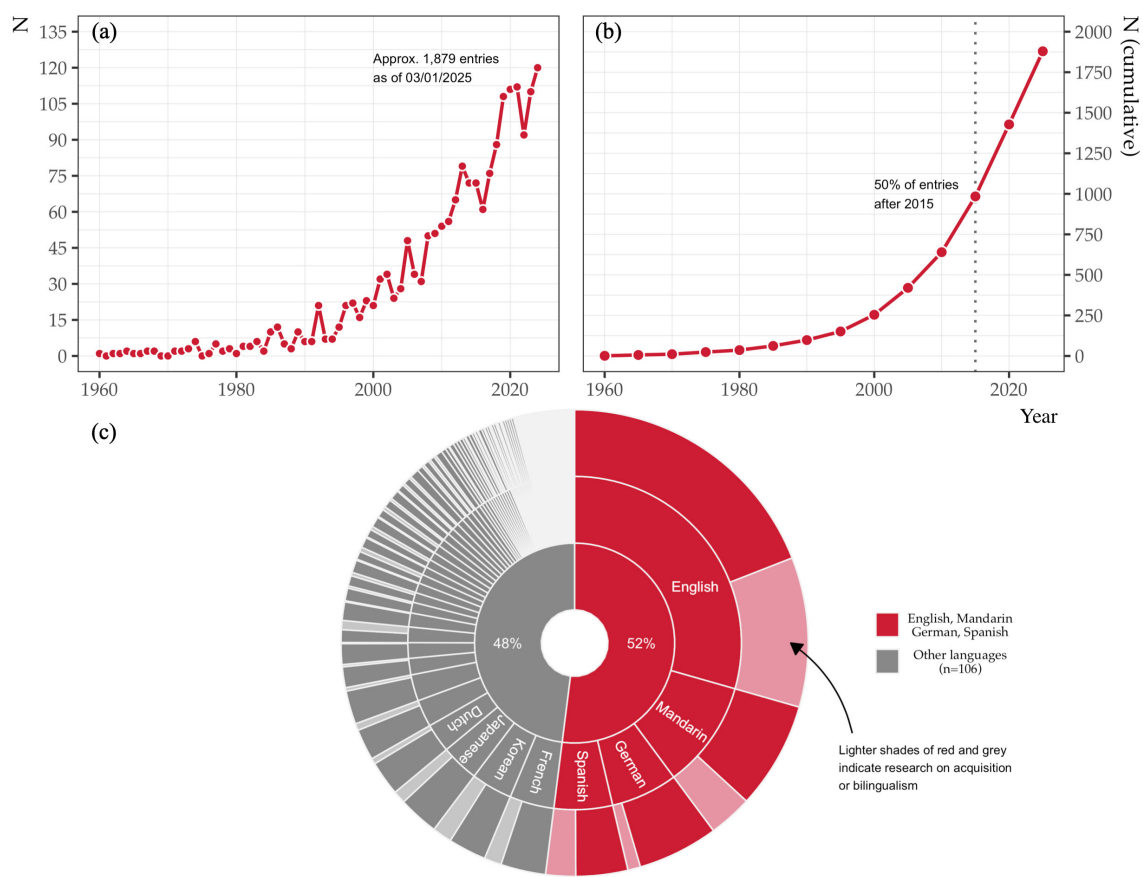


FIGURE 1
Number of articles per year (a) and cumulative totals in five-year intervals (b) featuring information structure terms in Scopus, 1960–2024. Proportion of Scopus articles on information structure by language and population, 1960–2024 (c).



OPEN ACCESS

EDITED BY

Tania L. Leal,
University of Arizona, United States

REVIEWED BY

Tiffany Judy,
Wake Forest University, United States
Tihana Kras,
University of Rijeka, Croatia

*CORRESPONDENCE

Cristóbal Lozano
✉ cristoballozano@ugr.es

RECEIVED 24 June 2023

ACCEPTED 18 October 2023

PUBLISHED 09 November 2023

CITATION

Lozano C and Quesada T (2023) What corpus data reveal about the Position of Antecedent Strategy: anaphora resolution in Spanish monolinguals and L1 English-L2 Spanish bilinguals.

Front. Psychol. 14:1246710.

doi: 10.3389/fpsyg.2023.1246710

COPYRIGHT

© 2023 Lozano and Quesada. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

What corpus data reveal about the Position of Antecedent Strategy: anaphora resolution in Spanish monolinguals and L1 English-L2 Spanish bilinguals

Cristóbal Lozano* and Teresa Quesada

Universidad de Granada, Granada, Spain

This study investigates the acquisition of anaphora resolution (AR) in Spanish as a second language (L2). According to the Position of Antecedent Strategy (PAS), in native Spanish null pronominal subjects are biased toward subject antecedents, whereas overt pronominal subjects show a “flexible” bias (typically toward non-subject but also toward subject antecedents). The PAS has been extensively investigated in experimental studies, though little is known about real production. We show how naturalistic production (corpus methods) can uncover crucial factors in the PAS that have not been explored in the experimental literature. We analyzed written samples from the CEDEL2 corpus: L1 English-L2 Spanish adult late-bilingual learners (intermediate, lower-advanced and upper-advanced proficiency levels) and a control group of adult Spanish monolinguals ($N = 75$ texts). Anaphors were manually annotated via a fine-grained, linguistically-motivated tagset in UAM Corpus Tool. Against traditional assumptions, our results reveal that (i) the PAS is not a privileged mechanism for resolving anaphora; (ii) it is more complex than assumed (in terms of the division of labor of anaphoric forms, their antecedents and the syntactic configuration in which they appear); (iii) the much-debated “flexible” bias of overt pronouns is apparent since they are hardly produced and are replaced by repeated NPs, which show a clear non-subject antecedent bias; (iv) at the syntax-discourse interface, the PAS is constrained by information structure in more complex ways than assumed: null pronouns mark topic continuity, whereas overtly realized referential expressions (overt REs: overt pronouns and NPs) mark topic shift. Learners show more difficulties with topic continuity (where they redundantly use overt pronouns) than with topic shift (where they normally disambiguate by using overtly realized REs), thus being more redundant than ambiguous, in line with the Pragmatic Principles Violation Hypothesis (PPVH) (Lozano, 2016). We finally argue that the insights from corpora should be implemented into experiments. The triangulation of corpus and experimental methods in bilingualism ultimately provides a clearer understanding of the phenomenon under investigation.

KEYWORDS

Spanish second language acquisition, anaphora resolution, position of antecedent strategy, learner corpora, pronominal subjects, CEDEL2 corpus

1. Introduction: anaphora resolution and the Position of Antecedent Strategy

Anaphora Resolution (AR) is a frequent and pervasive (though deceptively simple) mechanism found in all natural languages. Its acquisition represents a challenge for different types of bilinguals, including late sequential bilinguals like adult second language (L2) learners (Lozano, 2021a).

Anaphors like pronominal subjects refer to their antecedents in discourse. The ambiguous scenario in English (1) requires the resolution of the anaphor: *she* can refer to either antecedent (subject *Carmen* or object *Paola*). Null-subject languages like Spanish are anaphorically more complex since both null (\emptyset) and overt (*ella* “she”) pronouns can alternate in subject syntactic position, (2), and can refer to either antecedent. Despite this apparent ambiguity, our mental syntactic parser/processor has certain strategies to automatically resolve the anaphor.

- (1) Carmen_i greeted Paola_j while **she**_{i/j} was opening the door.
- (2) Carmen_i saludó a Paola_j mientras $\left\{ \begin{array}{c} \emptyset_{i/j} \\ \text{ella}_{i/j} \end{array} \right\}$ abría la puerta.

The Position of Antecedent Strategy (PAS),¹ originally formulated by Carminati (2002) for native Italian, resolves such ambiguity in intrasentential AR (subordinate-main clausal order). Carminati’s results from an offline sentence-interpretation task confirmed this trend: When asked about the interpretation of the second clause (e.g., *Who was in the United States?*), Italian monolinguals chose a subject antecedent (*Marta* 80.72%) with null pronouns in (3), but a non-subject antecedent (*Piera* 83.33%) with overt pronouns. Results from an online self-paced reading task (SPRT) confirmed this: null pronominals (\emptyset) take significantly shorter when referring to preverbal subjects (1,844 ms) than to postverbal objects (2,352 ms), whereas overt pronouns (*lei* “she”) take less time to non-subject (2,236 ms) than to subject (2,266 ms) antecedents.

- (3) Marta scriveva frequentemente a Piera quando $\left\{ \begin{array}{c} \emptyset \\ \text{lei}_i \end{array} \right\}$ era negli Stati Uniti.

“Marta wrote frequently to Piera when \emptyset /she was in the United States.”

The PAS is a syntactic/configurational parsing strategy with a clear division of labor: null pronouns are biased toward a preverbal subject antecedent whereas overt pronouns are biased toward a postverbal non-subject antecedent. Importantly, the PAS is also a syntax-discourse interface phenomenon due to the information status of the anaphor: null pronouns encode a continuation of the preceding subject (topic continuity), whereas overt pronouns mark a topic shift. This holds true in other null-subject languages like Spanish (Lozano, 2009, 2016, 2021a; Martín-Villena and Lozano, 2020), Moroccan Arabic (Bel and García-Alcaraz, 2015), Greek (Prentza and Tsimpli, 2013; Papadopoulou et al., 2015), Croatian (Kraš, 2008a,b), and Romanian (Geber, 2006), among other languages.

The PAS had been extensively investigated in diverse bilingual populations (adult and child L2 learners, heritage speakers, attriters) in different L1-L2 combinations, which has led to the proposal of key theories like Sorace’s (2011) Interface Hypothesis (IH), which predicts bilinguals to show limitations when simultaneously integrating syntactic and discursive information. Follow-up proposals, like Lozano’s (2016) Pragmatic Principles Violation Hypothesis (PPVH), locate the source limitations at a more pragmatic level (topic continuity vs. shift), as a result of the violation of pragmatic principles like Economy vs. Clarity.

Crucially, much of our understanding of AR in general and PAS in particular comes from experimental studies that (i) often report contradictory results, so it is still unclear how the PAS operates in native (and L2) Spanish, and (ii) repeatedly investigate similar anaphoric configuration (i.e., PAS). We argue that highly-contextualized, discourse-rich corpus production data can uncover many factors that have gone undetected in prior experimental studies and solves some of the unresolved PAS questions. Additionally, our developmental corpus data will also allow us to know how the PAS is acquired across proficiency in L1 English-L2 Spanish and whether very advanced learners can eventually acquire the pragmatic subtleties of PAS.

Carminati’s PAS was originally formulated for language processing (comprehension) and our aim is to put it to the test in language production (corpus data). In the psycholinguistic literature, it has long been acknowledged that “grammatical processing (or “parsing”) ... refers to the construction of structural representations for sentences, phrases and morphologically complex words in real-time language comprehension and production” (Clahsen and Felser, 2006, p. 564) and that “there may be a closer link between comprehension and production, in particular between parsing and syntactic encoding during production.” (Pickering and van Gompel, 2006, p. 487). In this line, Mac Donald (2013) empirically shows that “language production processes can provide insight into how language comprehension works” (p. 1) and concludes that “the availability of extensive language corpora in many languages permits comprehension researchers to examine the relationship between production patterns (in the corpus) and comprehension behavior” (p. 13). Additionally, it is widely acknowledged in the (bilingual) psycholinguistic literature (e.g., Fernández and Smith Cairns, 2011) that, during processing (parsing), two major processes take place: (i) structuring the incoming input into categories, and (ii) establishing appropriate dependency relations between such categories, which is particularly relevant when there is potential ambiguity (as is the case in PAS scenarios). AR in general and the PAS in particular are classic examples of dependency. Dependencies need to be established not only in comprehension (listener/reader’s perspective), but also in production since the speaker/writer needs to make sure that the anaphoric dependency s/he is producing is configurationally well established and structured (as is the case of PAS scenarios) to ensure that the listener/reader can interpret such dependency and therefore resolve the anaphor. Therefore, the use of production methods (corpora) can shed light on the PAS, as we do in this paper.

We next review the acquisition and processing of PAS in native and L2 Spanish based on experimental and corpus studies (section 1.1). In section 1.2 we present the research questions. The corpus methodology is discussed in section 2. Section 3 presents the results for each research question followed by a discussion, and section 4

¹ Also known as PAH (Position of Antecedent Hypothesis).

concludes with a general discussion/conclusion and future avenues of investigation.

1.1. The PAS in native and L2 Spanish

Overall, previous experimental native Spanish PAS findings show no clear division of labor as in native Italian: null pronouns select subject antecedents, but overt pronouns are “flexible” (non-subject and subject antecedents). Each experimental study is unique in terms of, e.g., the type of method/stimuli/design, which could explain the different results across studies. Consequently, we present a thorough review of each study to detect possible limitations that will be later implemented in our corpus study. Note that we review both offline and online PAS studies in adult Spanish monolinguals and adult L2 learners, thereby excluding other populations (see [Tables 1, 2](#) in the [online Supplementary material](#) for additional details).² Finally, no single corpus study has targeted PAS structures, so we review some corpus evidence on AR in general as their findings may shed light on PAS.

1.1.1. Offline experimental evidence

[Alonso-Ovalle et al. \(2002\)](#) administered a sentence interpretation task with intersentential PAS (4) to adult Peninsular Spanish monolinguals. Results from the comprehension question (Who is angry?) show a clear subject bias (*Juan* 73.2%) for null pronouns but a “flexible” behavior for overt pronouns (50.2% subject antecedent *Juan*, 49.8% non-subject antecedent *Pedro*), contra [Carminatti’s \(2002\)](#) original PAS formulation.

(4) Juan pegó a Pedro. $\left\{ \begin{matrix} \emptyset \\ \text{él} \end{matrix} \right\}$ está enfadado.

“Juan hit Pedro. (He) is angry.”

Adult Peninsular Spanish monolinguals (with knowledge of Catalan) were tested in an acceptability judgment continuation task, where the plausibility of the continuation sentence (*in italics*) was judged on a four-point scale ([Bel et al., 2016a](#)). Monolinguals judged main-subordinate clause order (5) vs. subordinate-main clause order (e.g., *Mientras Javier abandonaba a Pedro, se emborrachó. Pedro se emborrachó.*).

(5) Javier abandonó a Pedro mientras se emborrachaba. *Pedro se emborrachaba.*

“Javier abandoned Pedro while (he) was getting drunk. Pedro was getting drunk.”

When both clausal orders are analyzed together, null pronouns refer more to the subject (mean: 3.1) than the object (2.6), but overt pronouns refer to the object (3.2) more than to the subject (2.3). The same holds for *subordinate-main* order

(null: 3.55 subject, 2.25 object; overt: 3.25 object, 2.45 subject). This confirms Carminatti’s PAS. In *main-subordinate* order, results for the null pronoun were unexpected (null: 2.71 subject, 3.03 object; overt: object 3.01, subject 2.18). These unexpected monolingual finding led us to incorporate clausal order as a variable in our corpus-based study. The results for monolinguals were similar in [Bel and García-Alcaraz \(2015\)](#), who also included intermediate adult L1 Arabic-L2 Spanish learners in Morocco, both Moroccan Arabic and Spanish being null-subject languages with similar PAS behavior. Learners observed the PAS timidly in both clausal orders: (i) in *main-subordinate*, the null pronouns selected subjects (2.74 in main-subordinate, 2.64 in subordinate-main) slightly more than objects (2.54 and 2.34 respectively), but overt pronouns chose objects (2.81 and 2.63) more than subjects (2.16 and 2.40). In short, learners obey the PAS timidly, whereas Spanish/(Catalan) monolinguals do as well except for the main-subordinate condition, where null pronouns show the opposite behavior.

Jegerski and colleagues conducted a couple of PAS studies. First ([Jegerski et al., 2011](#)), they tested L1 English-L2 Spanish adult learners (intermediate, advanced) and adult Spanish monolinguals (from Spain and Latin America) in an ambiguous PAS sentence-interpretation task with null and overt pronouns (6).³

(6) Marta le escribía frecuentemente a Lorena cuando $\left\{ \begin{matrix} \emptyset \\ \text{ella} \end{matrix} \right\}$ estaba en los Estados Unidos.

“Marta wrote frequently to Lorena when (she) was in the United States”.

When asked about the anaphoric interpretation, monolinguals preferred to link null pronouns with subject antecedents (75%), as predicted by the PAS, but overt pronouns show again a “flexible” behavior (53% subject antecedents, 47% object antecedents). Advanced learners show a native-like tendency: null-subject 69%, and “flexible” overt pronoun behavior (56% subject antecedent, 44% object antecedent). Intermediates show a timid subject bias irrespective of the pronoun type (null-subject 66%, overt-subject 60%). In their second study, [Keating et al. \(2011\)](#) employed the same methodology and the same profiles of participants. Once again, Spanish monolinguals significantly preferred a null pronoun (74%) to an overt pronoun (54%) to refer to the subject. By contrast, the difference was not significant in advanced learners (60.15% null vs. 54.21% overt). Results from both studies indicate that overt pronouns show a “flexible” behavior by referring around 50% of the time to the subject and 50% to the object, both in native and L2 Spanish, a fact to which we will return in our study.

In a picture-verification task, [Clements and Domínguez \(2017\)](#) tested the PAS in adult monolinguals (mainly from Spain, some from Mexico) and advanced L1 English-L2 Spanish learners from the United Kingdom, who were presented with two pictures and a PAS sentence with(out) an overt pronoun, as in (6). They had to decide

² Note that in all the experimental studies under review, the stimuli always contain two potential antecedents (one in subject position, another in non-subject position). The advantage of using corpus data is that in natural production PAS structures typically contain more antecedents and in different syntactic positions (see sections 2.4 and 3.1).

³ The authors compared discourse-coordinating (*mientras* “while”) vs. -subordinating (*cuando* “when”/ *después de que* “after”/ *desde que* “since”) conjunctions. For brevity, we discuss the discourse-coordination results only.

whether the given sentence corresponded to one or the other picture (or both). Monolinguals preferred to link a null pronoun with a subject (77%) more than an object (12%) antecedent, whereas overt pronouns showed the opposite pattern (54% object, 27% subject), which supports Carminati's original PAS formulation, though note once again that the intuitions for overt pronouns are not as strong as those for null pronouns, a fact to which we will return in this paper. Unlike previous findings above, advanced learners observed the PAS in a native-like manner (null: subject 68%, object 21%; overt: object 63%, subject 23%).

Chamorro et al. (2016) asked adult monolinguals from Spain to rate null/overt pronoun PAS under four conditions: two forced antecedent-subject biases (singular subject, plural object (7a)), and two forced object-antecedent biases (plural subject, singular object (7b)). Monolinguals non-significantly rated the null pronoun to equally refer to the subject (3.72) and the object (3.61) antecedent, showing no clear subject bias of null pronouns, which runs against all the findings reviewed above. The overt pronoun significantly referred to the object (3.60) more than the subject (3.26) antecedent (though note the 3.26 vs. 3.60 ratings are not different enough given the 1–5 Likert rating scale).

- (7) a. La madre_i saludó a las chicas_j cuando $\left\{ \begin{matrix} \emptyset_i \\ \text{ella}_i \end{matrix} \right\}$ cruzaba una calle con mucho tráfico.
 b. Las madres_j saludaron a la chica_i cuando $\left\{ \begin{matrix} \emptyset_j \\ \text{ella}_j \end{matrix} \right\}$ cruzaba una calle con mucho tráfico.

“The mother (s) greeted the girl (s) when (she) was crossing a street with lots of traffic.”

In a picture selection task, Martín-Villena (2023) tested conjunction type (when vs. while) in Peninsular Spanish monolinguals in sentences like (6). Subject-antecedent preferences with conjunction *cuando* “when” were higher for null (67%) than overt (23%) pronouns as well as with *mientras* “while” (null: 80%; overt: 30%). This confirms PAS preferences for subject antecedents but shows that null-subject bias was somewhat stronger with *mientras* “while” than with *cuando* “when”.

1.1.2. Online experimental evidence

All online experiments to date have used SPRT, which measure reading times (RTs) in milliseconds (ms). Filiaci (2010) was the first online study to test PAS in Peninsular Spanish monolinguals. In intrasentential subordinate-main clauses, (8), the semantics of the main clause forced the anaphor toward the subject (8a) or the object (8b) antecedent. RTs of the main clause with a null pronoun were significantly faster when biasing toward the subject (1,998 ms) than the object (2,319 ms) antecedent, as predicted by Carminati's PAS, but with an overt pronoun, RTs were faster when biasing toward the object (2,389 ms) than the subject (2,507 ms) (but differences were non-significant, which reflects again the “flexible” behavior of Spanish overt pronouns). These results were later published (Filiaci et al., 2014) as experiment 1. Experiment 2 stimuli were the same as in experiment 1 but RT analyses were conducted at different phrasal regions (separated by slashes “/” in (9)). Overall, results replicated those found in experiment 1, thus confirming the “flexibility” of overt pronouns in Spanish when compared to Italian.

- (8) a. Cuando Ana_i visitó a María_j en en el hospital, $\left\{ \begin{matrix} \emptyset_i \\ \text{ella}_i \end{matrix} \right\}$ le llevó un ramo de rosas.
 b. Cuando Ana_i visitó a María_j en en el hospital, $\left\{ \begin{matrix} \emptyset_j \\ \text{ella}_j \end{matrix} \right\}$ ya estaba fuera de peligro.

“When Ana visited Mary in the hospital, (she) {brought her a bunch of roses | was already out of danger.}”

- (9) Cuando / Ana / visitó / a María / en en el hospital, / $\left\{ \begin{matrix} \emptyset_i \\ \text{ella}_i \end{matrix} \right\}$ / le llevó / un ramo / de rosas.

Gelormini-Lezama and Almor (2011) tested intersentential PAS with adult Argentinian Spanish monolinguals. Sentences also included repeated names (RNs) (e.g., Juan “John”), (10). The object clitic (*la* “her”) forces the null pronoun toward a subject (10a) or object (10b) antecedent reading. With forced subject antecedents, RTs for null-pronoun sentences (1,812 ms) were faster than overt-pronoun sentences (2,264), but the opposite was true when with forced object antecedents (null 2,412, overt 2,157). This clearly confirms Carminati's PAS prediction. Interestingly, RNs were read equally fast irrespective of their antecedent (2080 subject, 2055 object) and their RTs did not significantly differ from sentences containing overt pronouns but did significantly differ from sentences containing null pronouns (subject: null < RN; object: null > RN), which suggests that NPs may play a role in object-antecedent selection in AR in native Spanish, a fact to which we will return in our corpus analysis.

- (10) a. Juan_i se encontró con María_j. $\left\{ \begin{matrix} \emptyset_i \\ \text{Él}_i \\ \text{Juan}_i \end{matrix} \right\}$ la_j vio triste.
 b. María_j se encontró con Juan_i. $\left\{ \begin{matrix} \emptyset_j \\ \text{Él}_j \\ \text{Juan}_j \end{matrix} \right\}$ la_i vio triste.

“{John found Mary | Mary found John}. \emptyset /He/John found her sad.”

Another study (Bel et al., 2016b) tested adult Peninsular Spanish monolinguals in intrasentential (main-subordinate order) PAS, (11), presented in a word-by-word, non-cumulative fashion. The ambiguous anaphor is resolved postverbally via world knowledge: *violin* forces a subject antecedent (musician), whereas *casco* “helmet” forces an object antecedent (firefighter).

- (11) El músico_i saluda al bombero_j mientras $\left\{ \begin{matrix} \emptyset_i \\ \text{él}_i \end{matrix} \right\}$ lleva $\left\{ \begin{matrix} \text{un violín} \\ \text{un casco} \end{matrix} \right\}$ en la mochila.

“The musician greets the fireman while (he) carries {a violin | a helmet} in his backpack.”

Null pronouns were read significantly faster with a subject-antecedent bias (798 ms) than an object-antecedent bias (887 ms) at the NP object region (*un violín/un casco*), but not at the locative PP region (*en la mochila*) (1,143 vs. 1,453 ms). By contrast, overt pronouns were read significantly faster with an object-antecedent

bias (1,308 ms) than with a subject-antecedent bias (1,402 ms) at the PP region, but not at the object region (884 vs. 887 ms). Findings are in line with Carminatti's PAS prediction, though note that (i) RT differences⁴ for overt pronouns (170 ms) are smaller than for null pronouns (399 ms), which suggests again a rather "flexible" antecedent bias for overt pronouns; (ii) RT differences are more observable in some regions than in others, which suggests that these stimuli are not straightforwardly parsed probably due to the complex disambiguation mechanism. Further results from adult L1 Arabic-L2 Spanish and L1 English-L2 Spanish learners at three proficiency levels (intermediate, upper intermediate, high) revealed that the advanced learners can eventually parse PAS structures in a native-like fashion, irrespective of their L1 (a (non)null-subject language like English or Arabic).

Intrasentential (subordinate-main order) PAS was investigated in adult Mexican Spanish monolinguals (clause-by-clause presentation) (Keating et al., 2016). The ambiguous anaphor is resolved postverbally via world knowledge: *su culpabilidad* "his guilt" forces a subject antecedent (*el sospechoso* "the suspect") in (12a), but an object antecedent in (12b). Null-pronoun clauses were read significantly faster with subject (2,186 ms) than with object (2,447 ms) antecedents. By contrast, overt-pronoun sentences were read faster with object (2,456 ms) than with subject (2,605 ms) antecedents. These results confirm Carminatti's PAS but note that if we calculate the RT differences,⁵ the mathematical difference is smaller for overt pronouns (194 ms) than for null pronouns (261), which suggests again a certain "flexibility" for overt pronouns.

- (12) a. Después de que el sospechoso_i habló con el policía_i, $\left\{ \begin{matrix} \emptyset_i \\ \text{él}_i \end{matrix} \right\}$
admitió su culpabilidad.
- b. Después de que el policía_j habló con el sospechoso_j, $\left\{ \begin{matrix} \emptyset_j \\ \text{él}_j \end{matrix} \right\}$
admitió su culpabilidad.

"{After the suspect talked to the policeman | After the policeman talked to the suspect}, (he) admitted his guilt."

In a SPRT, Martín-Villena (2023) used the same stimuli as in the offline experiment above. Results showed differences depending on the region analyzed. In the subordinate clause segment, null pronouns showed an unclear bias (Subject: 1,383 ms; Object: 1,372 ms), but overt pronouns exhibited a clear object bias (Subject: 2,129 ms; Object 1,940 ms). Interestingly, in the comprehension question segment, null pronouns showed a subject bias (Subject: 946 ms; Object 1,051 ms), but overt pronouns showed an object bias (Subject 1,242 ms; Object: 1,037 ms), as predicted by the PAS.

1.1.3. Summary of the experimental evidence: native Spanish

The native Spanish PAS results from the experimental studies are often contradictory. This could be due to multiple factors (many of

which were taken into account in our corpus study), e.g.: the different varieties of the monolinguals of Spanish; the PAS configuration (intersentential vs. intrasentential) and the clausal order (main-subordinate vs. subordinate-main); and the different formats (and presentation types) of the offline and online experimental methods, among others.

A visual summary of *offline* PAS biases in native Spanish (Figure 1) suggests that the original PAS formulation for native Italian is not fully operative in native Spanish: Whereas null pronouns clearly select a subject antecedent (69%~87% range), as predicted by Carminatti's PAS, overt pronouns show a "flexible" preference by often selecting an object antecedent around half of the time (50%~65% range), which implies that the rest of the time they select a subject antecedent. In *online* experiments, null-subject sentences are read significantly faster with forced subject than with forced object antecedents, whereas overt-subject sentences are read faster with forced object than subject antecedents, as predicted by PAS, though note that the subject vs. object RT differences are usually weaker with overt pronouns than with null pronouns, which again suggests a mild "flexibility" of overt pronouns. The offline and online native Spanish findings thus suggest that, whereas null pronouns have a strong subject bias, overt pronouns are less clear-cut (i.e., more "flexible") in their choice of antecedent. We will argue that such flexibility is more apparent than real, as our corpus data will reveal.

1.1.4. Corpus evidence

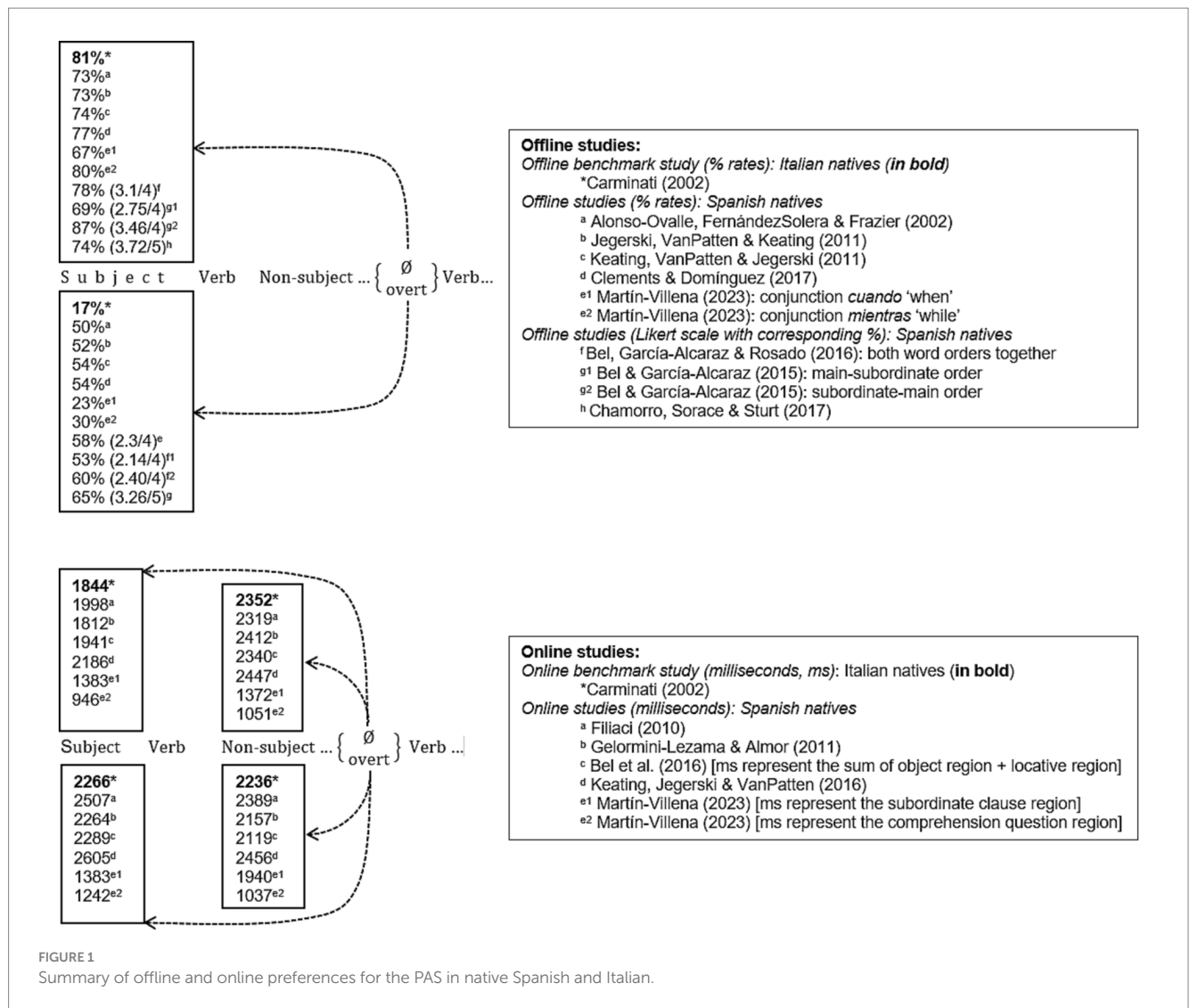
To our knowledge, there is no corpus-based study targeting specifically the PAS in adult Spanish monolinguals/learners. At best, there is some indirect PAS evidence since the corpus studies to be reviewed analyzed multiple types of AR scenarios (including PAS), so it is unclear to what extent their findings can extrapolate to specific PAS scenarios.

The experimental study reviewed above (Bel et al., 2016a) presents additional evidence from a written and spoken production task by Peninsular Spanish monolinguals. The researchers analyzed different types of AR scenarios, including PAS-like scenarios. Null pronouns clearly biased toward subject antecedents (77.27%), while overt pronouns showed a less clear-cut antecedent bias (subject: 42.86%; non-subject: 57.14%). Moroccan Arabic/Spanish early bilinguals' null pronouns clearly biased toward subject antecedents (70.19%), while their overt pronouns biased toward both non-subject (35.71%) and subject (64.29%) antecedents. Overt pronouns reflect again the already reported "flexibility". Importantly, this study (i) does not report the production of NP anaphors, which are crucial for our understanding of AR in general and the PAS in particular, as we will later show in this paper; (ii) analyses both singular and plural anaphoric forms together, though corpus data has shown that only 3rd singular anaphors are problematic for learners (Lozano, 2009, 2016); and (iii) presents data from teenage Spanish monolinguals⁶ and early bilinguals, so the evidence about how the PAS operates in adult monolinguals and L2 Spanish is rather indirect. In a follow-up study, García-Alcaraz and Bel (2019) used the same task and coding criteria. This time, the Spanish monolinguals were university students, and the L1 Moroccan

⁴ RT differences = subject antecedent (RT of object region + RT of PP region) – object antecedent (RT of object region + RT of PP region).

⁵ RT differences = subject antecedent (RT of main clause) – object antecedent (RT of main clause).

⁶ The discursive/pragmatic properties of AR are not fully acquired until around 15 years of age (Shin and Smith Cairns, 2012), so evidence from these teenage monolinguals should be taken cautiously.



Arabic-L2 Spanish learners were teenage sequential bilinguals. Results suggest that both monolinguals and bilinguals produce null pronominal subjects to mark topic continuity around 2/3 of the time and topic shift around 1/3. Regarding overt pronouns, their production was very low (4 tokens or less depending on the configuration), which is not very informative. In short, while suggestive, these findings do not fully inform about PAS scenarios in either native or L2 adult Spanish.

A series of corpus studies (Lozano, 2009, 2016; Martín-Villena and Lozano, 2020) targeted AR scenarios with subject anaphoric forms (null/overt pronouns, as well as NPs). Results from adult Peninsular Spanish monolinguals reveal some consistent findings across studies: whereas null pronouns clearly encode topic continuity, it is NPs that encode topic shift more often than overt pronouns do, particularly when there are several potential antecedents in discourse. L1 English-L2 Spanish learners do not typically show problems in topic-shift contexts (as they use overt forms to avoid ambiguity) but are redundant in topic-continuity contexts (as they overuse overt pronouns). These findings are captured by the Pragmatic Principles Violation Hypothesis (PPVH) (Lozano, 2016), which postulates differential effects at the syntax-discourse interface with AR: learners

obey the pragmatic Principle of Clarity as they use full anaphoric forms in cases of ambiguity, but they are lax with the Principle of Economy, as they redundantly produce overt anaphoric forms when not required in topic continuity, though can be modulated by the amount of potential antecedents. In short, learners are more redundant than ambiguous. We will get back to the PPVH when discussing our results.

To summarize, the corpus-based findings are clearly insufficient since they: (i) do not specifically target PAS scenarios but rather conflate different types of AR scenarios in their analyses; (ii) some of them do not consider the role of subject NPs as an anaphoric form in its own right. This, coupled by certain limitations in the experimental studies, motivated the formulation of our research questions with a view to answering some unresolved issues in the production of PAS in native and L2 Spanish.

1.2. The current study: research questions and hypotheses

The bulk of experimental studies on AR have investigated the PAS with two potential antecedents (subject/non-subject) and two

anaphoric forms (overt/null pronominal subject) in either inter- or intra-sentential configurations. So, what we know about the PAS comes mostly from a series of similarly-designed experiments that do not question whether (i) the PAS may represent an oversimplified way of resolving anaphora in native (and L2) Spanish; (ii) PAS scenarios may be more complex than traditionally assumed (i.e., they can contain more than two antecedents in other syntactic positions); (iii) the antecedents may be realized by other forms other than null/overt pronouns (i.e., NPs for example). Unlike experiments, corpus data can shed light on these questions since they contain natural language production (where AR configurations are neither controlled nor constrained) and offer contextually rich scenarios with anaphors and antecedents embedded in their entire discourse. Unlike experiments, corpus data can shed light on these questions since they (i) contain natural language production where AR configurations are neither controlled nor constrained; (ii) offer contextually rich scenarios with anaphors and antecedents embedded in their entire discourse. This led to RQ1a and RQ1b.

RQ1a (Prototypicality of PAS): Is the PAS a prototypical way of resolving anaphora in native (and in non-native) Spanish, as implicitly assumed in the literature?

H1a: The PAS is but one of many possible mechanisms for resolving anaphors in native and non-native Spanish.

RQ1b (Complexity of PAS): Can the standard PAS configuration (subject/non-subject antecedent; null/overt pronominal subject anaphor) be more complex than assumed in the literature?

H1b: Corpus data will reveal that the PAS is richer than standardly assumed, in terms of antecedent configurations, syntactic possibilities and range of anaphoric forms.

Experimental PAS studies have typically restricted their focus to two anaphoric forms (overt/null pronominal subjects). Corpus studies have reported the use of other anaphoric forms (e.g., repeated Ns and NPs) in several AR scenarios, so NPs may be also possible Referential Expression (RE) forms in PAS.⁷

RQ2 (RE forms in discourse): Apart from null/overt pronominal subjects, are other RE forms possible in native and L2 Spanish PAS?

H2: In line with corpus findings on AR in general, we predict for PAS (i) null pronouns to be abundant due to the null-subject nature of Spanish; (ii) overt pronouns to be infrequent and, (ii)

importantly, NPs to be more frequent than overt pronouns. The range of REs in PAS will therefore include null/overt pronominal anaphors and NPs (used with an anaphoric value).

Experimental studies report Spanish null pronouns to bias toward a preverbal subject antecedent, whereas overt pronouns show a more “flexible” behavior. This contrasts with native Italian where overt/null pronouns show a clear division of labor. Additionally, experimental studies have not typically included NPs as a possible RE form.

RQ3 (Division of labor): Regarding the division of labor in native and L2 Spanish, will the “flexible” behavior of overt pronouns be better accounted for if NPs are also included as a possible type of RE?

H3: Null pronouns will be clearly biased toward a subject antecedent, as previously reported, whereas overtly realized REs (i.e., overt pronouns and NPs together) will be clearly biased toward non-subject antecedents. Learners will show growing sensitivity to such division of labor as proficiency increases, but native-like ultimate attainment is not expected for upper-advanced learners since the PAS is constrained at the syntax-discourse interface (cf. RQ4 below), which is a problematic area for L2 learners (Lozano, 2021a for an overview).

The implicit assumption in the experimental literature is that purely configurational factors (null→subject vs. overt→non-subject) overlap with discursive information-status factors (null→topic continuity vs. overt→topic shift). RQ4/H4 (when contrasted to RQ3/H3) will determine the extent to which the overlap assumption is correct. This motivates theoretical questions having to do with likely deficits at the syntax-discourse interface.

RQ4 (Syntax-discourse interface): Will syntactic configuration overlap with information status in PAS configurations and, if so, will learners be eventually (un) able to acquire this syntax-discourse phenomenon?

H4: Syntactic configuration will overlap with information status and NPs will play a role (null→subject/topic continuity; overt & NP→non-subject/topic shift). Learners will show an increasing trend toward the native norm, yet the syntax-discourse properties of the PAS will not be fully acquired, as predicted by models like the IH and the PPVH.

Despite English being a non-null subject language, corpus data (Quesada and Lozano, 2020) have shown that English monolinguals allow null pronouns in very specific contexts: topic continuity and coordination at around 77% (e.g., *Lucy_i walked for an hour and Ø_i had a picnic*), but never in non-coordinate configurations. So, it could be argued that L2 Spanish learners’ production of null pronouns in topic continuity could be due to L1 transfer rather than actual acquisition, which leads to the following exploratory research question.

RQ5 (Cross-linguistic influence): Will L2 Spanish learners’ distribution of null pronouns be a reflection of their allowance

⁷ We incorporate NPs and repeated proper Ns as type of anaphoric form, hence we use the wider term Referential Expressions (REs) to include all forms (overt/null pronouns, NPs, repeated Ns), instead of the more restrictive term anaphoric forms.

in their L1 English (topic continuity and coordination) or will it be a reflection of acquisition at the syntax-discourse interface? It may be the case that learners transfer in initial stages but progressively acquire the discursive distribution of null pronouns.

H5: (Transfer account)

If L2 Spanish learners are transferring from their L1 English, null subjects will be produced mainly where they are allowed in English (topic continuity with coordination) and not where they are not allowed (topic continuity with non-coordination).

(Non-transfer account, i.e., acquisition account)

If they are rather sensitive to the pragmatics of null pronouns in Spanish, null subjects will be produced where they are allowed in native Spanish, i.e., across the board (both in coordination and non-coordination).

Previous PAS experimental studies are often contradictory depending on the sentential configuration: inter- vs. intra-sentential; main-subordinate vs. subordinate-main orders (*cf.* the tables in the [online Supplementary material](#)). RQ6 is an exploratory question to explore whether the sentential PAS configuration modulates the choice of RE in naturalistic corpus production.

RQ6 (Sentential configurations): In which sentential configurations (intra- vs. inter-sentential) will PAS structures be more frequent in naturalistic corpus production? Which PAS clausal order (main-subordinate vs. subordinate-main) is prototypical? Will learners' production ultimately approach to/deviate from Spanish monolinguals?

2. Method

2.1. Corpus: CEDEL2

Corpus Escrito del Español L2 (CEDEL2) ([Lozano, 2022](#)) is a multi-L1 corpus of L2 Spanish learners coming from 11 different L1 backgrounds, plus a Spanish monolingual control subcorpus. CEDEL2 (version 2) currently holds 1,105,936 words, 4,399 participants, and 14 task topics. It is freely available/downloadable at <http://cedel2.learnercorpora.com>.

Data are collected via online forms⁸ and participants complete three forms: (i) linguistic background; (ii) standardized placement test (just for learners) ([University of Wisconsin, 1998](#)); and (iii) written/spoken text.

2.2. Sample

We selected an L1 English-L2 Spanish (plus a comparable Spanish monolingual control) sample ([Table 1](#)) based on the following criteria: (i) the participant's age range was 18 ~ 40, since Working Memory, which may affect AR, appears to decay after the age of 40 ([Bel et al., 2016b](#)); (ii) learners' proficiency-level range was intermediate~advanced; and (iii) only two composition titles were targeted (*cf.* 2.3 below).⁹ Two hundred two texts met these criteria but we finally selected those that had at least one instance of a PAS ($N=75$). We originally departed from two intermediate groups: lower intermediates (placement score: 21 ~ 28 raw score, 49% ~ 65%) and upper intermediates (29 ~ 35, 67 ~ 81%). Since they did not significantly differ in our analyses, we decided to analyze both groups as a single group of intermediates to simplify the between-group statistical analyses and interpretations. Learners had an equivalent age of exposure (AoE) to L2 Spanish and their length of instruction (LoI) in

⁸ <http://learnercorpora.com>

⁹ Only monolinguals from Spain were chosen since in certain varieties (Mexican, Caribbean, Puerto Rican), overt pronouns mark topic continuity ([Flores-Ferrán, 2004](#)).

TABLE 1 Texts and participants' bio-data.

Group	Intermediate		Lower advanced	Upper advanced	Monolinguals
	Lower	Upper			
Placement raw score (0–43)	21 ~ 28	29 ~ 35	36 ~ 40	41 ~ 43	–
Equivalent percentage (0–100%)	49 ~ 65%	66 ~ 81%	82 ~ 94%	95 ~ 100%	–
Texts that met criteria	69		37	13	103
Texts analyzed	21		19	8	27
Mean age	20.8		21.3	25.5	25.6
Mean proficiency	69.6%		86.8%	96.7%	–
AoE (years)	14		14.5	12.2	–
LoI (years)	5.3		5.5	8.8	–
LoSA (months)	7.4		11.1	12.3	–

Spanish and length of stay abroad (LoSA) in a Spanish-speaking country increased with proficiency.

2.3. Task

We selected two task titles (*Talk about a famous person* and *Summarize a film you have recently seen*), since they are narratives that contain (i) abundant [+human] 3rd person antecedent-anaphor chains; and (ii) PAS constructions, which were more frequent in the second task than in the first task and which offered different characters in discourse suitable for the topic continuity/shift purpose of this study.

2.4. Corpus annotation and tagset

The corpus sample was manually annotated (i.e., tagged) with UAM Corpus Tool (O'Donnell, 2009), version 6.2j (February 2023).¹⁰ We firstly tagged each text to indicate the group category (intermediate, lower advanced, upper advanced, and monolingual), which allows between-group comparisons for the same linguistic feature, as will

be explained below. We designed another tagset to count the frequency of two AR scenarios (PAS vs. other AR). Each RE in subject position was assigned either the *PAS* tag (when the RE was preceded by a subject/non-subject antecedents) or *other* (when the RE was preceded by an AR scenario other than PAS). Figure 2 shows the fine-grained, linguistically-informed tagset to annotate PAS.¹¹ It allows for multiple and intricate statistical analyses among tags, as will become obvious later. It is inspired by previous corpus studies on AR (Lozano, 2016; Quesada and Lozano, 2020), although we introduced new features.

Every 3rd person human subject that followed the syntactic configuration of the PAS was manually tagged. First, the **PAS-type** system included: (i) *standard* PAS with two antecedents, as in (13), and (ii) *complex* PAS with more than two antecedents, as in (14a-c). For example, the tag used to annotate complex PAS in (14a) is *s1_nons2_nons3*, which indicates we have 3 potential antecedents: the first one is in subject position (s1) and the other two in non-subject position realized via a complex NP: PP (nons3) within an NP (nons2). In (14b), the tag *s1_nons2andnons3* indicates that there is a singular antecedent in subject position (s1, which

11 The original tagging scheme included a richer tagset with more tags that are not analyzed in this study due to space limitations –see Quesada (2021) for details.

10 <http://www.corpustool.com>

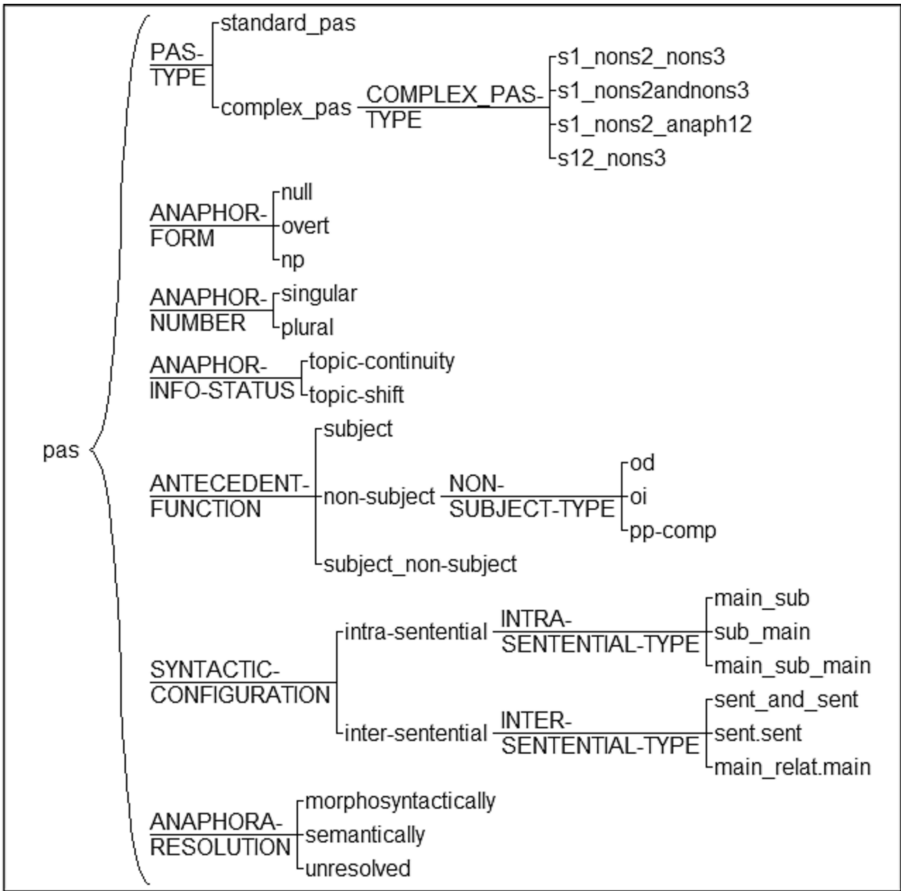


FIGURE 2
PAS annotation tagset.

happens to be a null pronouns) followed by two NP coordinated antecedents in non-subject position (s2&s3) embedded within a PP. Notice that, due to the complexity of the antecedents' region, the anaphor is a complex NP for disambiguation purposes. Other complex PAS contained plural REs, as in (14c), but we excluded them from our current analysis since it has been shown that the truly problematic cases of AR are 3rd person singular and not plural (Lozano, 2009).

(13) Standard PAS:

Naaven_i se ha enamorado de Tiana_j y Ø_i quiere pedirle matrimonio. [Monolingual: ES_WR_24_3_IZG.txt].¹²

"Naaven_i has fallen in love with Tiana_j and Ø_i wants to propose to her".

(14) Complex PAS:

a. *La chica_i se enamora del amante_j de su madre_k hasta que al final Ø_i acaba teniendo ...* [Monolingual: ES_WR_30_3_JVM].

"The girl_i falls in love with the lover_j of her mother_k until Ø_i ends up having..."

b. *Pero el principal problema que Ø_i tenía era que Ø_i sufría un maltrato constante por parte de su madre_j y del novio_k de ésta. El novio_k de la madre_j había...* [Monolingual: ES_WR_31_3_EAC]

"But the main problem Ø_i had was that Ø_i was abused by her mother_j and the boyfriend_k of her_j. The boyfriend_k of the mother_j had..."

c. *Ø_{ij} Juntos tendrán que huir de Dr. Facilier_k a los pantanos, dnde Ø_{ij} se encuentran...* [Monolingual: ES_WR_24_3_IZG]

"Ø_{ij} Together will have to escape from Dr. Facilier_k to the swamps, where Ø_{ij} meet ..."

The **anaphor-form** system includes the RE form (null/overt pronouns and NPs) in subject position, as shown in bold in (15). The **anaphor-number** system includes the RE number (singular/plural), which served us to exclude plural REs in the analyses, as justified above.

(15) ... **el protagonista_i** de la película se enamora de la chica_i y **ella_i**, le pide por favor que Ø_i deje el negocio ... [Monolingual: ES_WR_23_3_EM].

"...the **main character_i** of the film falls in love with the girl_i and **she_i** asks him_i that Ø_i leaves the business ..."

The **information-status** system comprises topic-continuity and topic-shift contexts, as in (16a-b) respectively. The

antecedent-function system included subject antecedent, non-subject antecedent, and subject/non-subject antecedent (for cases of complex PAS). This system allowed us to detect PAS scenarios with subject-antecedent biases, as in (16a), or non-subject antecedent biases, as in (16b).

(16) a. *Un periodista_i investiga la desaparición de una rica heredera_j, hace cuarenta años. Para ello, Ø_i cuenta con...*

[Monolingual: ES_WR_24_3_AW].

"A journalist_i investigates the disappearance of a rich heiress_j, 40 years ago. To do so, Ø_i relies on..."

b. *Bella_i se da cuenta de que Jacob_j está enamorado de ella, y **ella_i** también un poco de él* [Monolingual: ES_WR_21_3_ICH].

"Bella_i realizes that Jacob_j is in love with her_i and **she_i** is also in love with him_j".

In the **syntactic-configuration**, we tagged the type of intra-sentential and inter-sentential configurations, e.g., topic-continuity and coordination in (16b) and topic continuity and non-coordination, which can be of different types, e.g., subordination in (17) or new sentence in (18).

(17) ... un padre_i trata por todos los medios de llevar a su hijo_j de diez años hasta el mar, donde Ø_i espera encontrar...

[Monolingual: ES_WR_22_3_AFL].

"...a father_i tries by all means to take his ten-year-old son_j to the sea, where Ø_i hopes to find ..."

(18) ... y Ø_i llega a cortar_i un dedo de un hachazo. Después Ø_i intenta matar a George_k... [Monolingual: ES_WR_28_3_MAAO].

"... and Ø_i cuts off her_j finger with an axe. Later Ø_i tries to kill George_k..."

Finally, the **anaphora-resolution** system indicates the type of resolution: via morphosyntax or semantics. In this paper, we analyzed only the PAS that was morphosyntactically resolved. In order to avoid skewing our results, we excluded PAS that was semantically resolved (i.e., null pronouns in topic-shift scenarios like (19), which are ultimately resolved via directive verbs).

(19) *Ella_i le_j pide que Ø_j espere a...* [Monolingual: ES_WR_26_3_MPVI].

"She_i asks him_j that Ø_j waits for her_i to..."

2.5. Analysis

UAM Corpus Tool has an in-built statistical analysis software. Between-group (or between-system/tag) comparisons are based on the tags' raw frequencies and statistical contrasts are chi square (χ^2) tests, accompanied by their significance level (p) and their effect size (Cohen's h).

¹² After each corpus example, we provide in square brackets the filename from the CEDEL2 corpus (<http://cedel2.learnercorpora.com>).

Based on the linguistically-motivated tagging scheme (Figure 2), UAM Corpus Tool allows for multiple and sophisticated statistical contrasts between the different groups and the (sub) nodes and terminal nodes of the tagset. These contrasts were motivated by the linguistically-informed hypotheses from section 1.2. Following statistical recommendations for corpus data (Egbert et al., 2020), we purposely decided to use the χ^2 statistical contrasts provided by the software rather than submitting the data to more sophisticated statistical analyses (which involve transforming the data and abstracting away from the linguistic facts and interpretations):

“the most appropriate method for the task at hand should not be the most sophisticated method ... Instead, we should always strive to choose minimally sufficient statistical methods, meaning that we should choose tests that are no more nor less sophisticated than the study design requires. The reason for this is twofold: (1) all descriptive and inferential statistical tests force us to abstract away from language to some extent and (2) there is often an inverse relationship between the level of sophistication of the method and the linguistic interpretability of the results.” (Egbert et al., 2020, p. 40)

3. Results and discussion

We next present and discuss the results for each research question. We leave the general discussion for section 4.

3.1. RQ1/H1: frequency of PAS scenarios in natural language production

In Figure 3, PAS scenarios (gray bars) were compared against other types of AR scenarios (black bars). Spanish monolinguals

resolve anaphora via scenarios (68.2%, i.e., 296 REs out of a total of 434 tagged REs) other than standard PAS (21.2%) or complex PAS (10.6%). Thus, standard PAS only amounts to around 1/5th of the total possible AR scenarios. Learners show a similar pattern to monolinguals across all proficiency levels, though only the upper advanced group shows native-like behavior (standard and complex PAS $\chi^2=2.16$, $p=0.1419$ n.s., $h=0.204$; other scenarios $\chi^2=0.15$, $p=0.6964$ n.s., $h=0.030$). The lower-level learner groups significantly differ from Spanish monolinguals in other scenarios but not in standard and complex PAS scenarios (intermediates vs. monolinguals: standard and complex PAS $\chi^2=0.49$, $p=0.4848$ n.s., $h=0.087$, other scenarios $\chi^2=8.80$, $p=0.0030$, $h=0.193$; lower-advanced vs. monolinguals: standard and complex PAS $\chi^2=0.66$, $p=0.4180$ n.s., $h=0.106$, other scenarios $\chi^2=12.44$, $p=0.0004$, $h=0.235$).

Our findings support *H1a* (PAS represents one of the many possible mechanisms of AR in native and non-native Spanish) and *H1b* (PAS can contain more complex configurations than those traditionally reported in the literature). Corpus data reveal that the traditional assumption of standard PAS as a prototypical strategy to resolve anaphora has been overestimated in the experimental literature.

3.2. RQ2/H2: overall use of REs in PAS scenarios

RQ2 explores the different RE forms in PAS scenarios, independently from the factors that constrain their choice. Spanish monolinguals produced mostly null pronominal subjects (66.1%), followed by NPs (23.2%) and overt pronominal subjects (10.7%) (Figure 4). Learners show a tendency toward the native norm as proficiency increases, yet only upper-advanced learners (57.9% null, 26.3 overt, 15.8% NP) show a rather similar and non-significant pattern to the Spanish monolinguals (null pronouns: $\chi^2=1.30$,

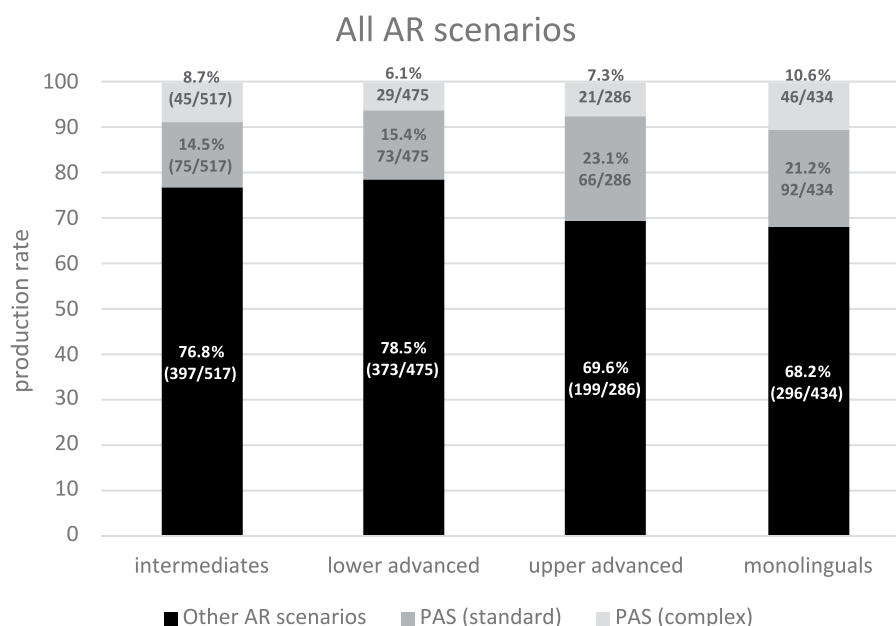


FIGURE 3
AR scenarios by group.

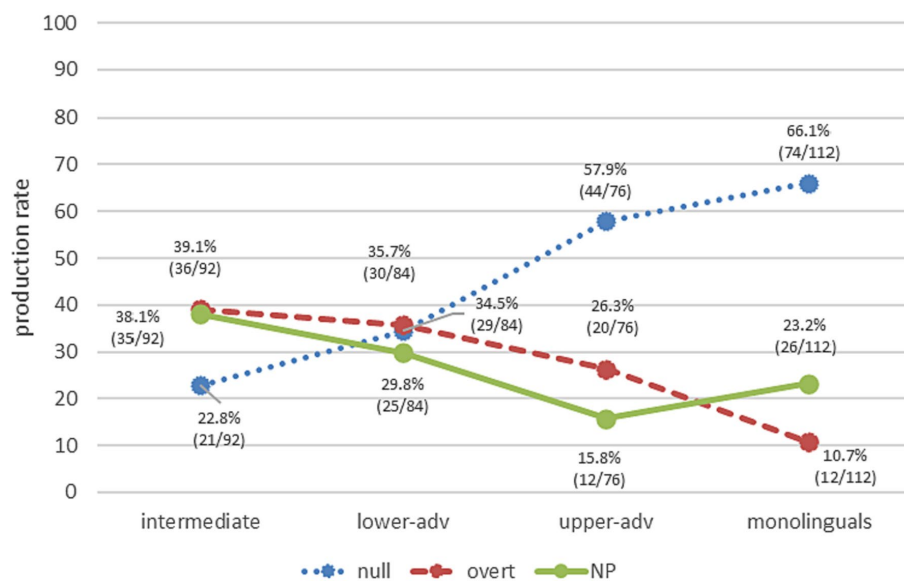


FIGURE 4
Overall production of REs across groups.

$p=0.2551$, $h=0.169$; NPs: $\chi^2=1.55$, $p=0.2135$, $h=0.188$), though a significant difference for overt pronouns ($\chi^2=7.80$, $p=0.0052$, $h=0.410$). The lower-advanced group shows similar proportions for all three RE forms (34.5% null, 35.7% overt, 29.8% NP), which significantly differ from monolinguals for null ($\chi^2=19.16$, $p<0.001$, $h=0.642$) and overt ($\chi^2=17.82$, $p<0.001$, $h=0.614$), but are non-significant for NPs ($\chi^2=1.07$, $p=0.3012$, $h=0.149$). Intermediates produce mainly overt REs (overt pronouns 39.1%; NPs 38.1%) and some null pronouns (22.8%), with the three RE production rates being significantly different from monolinguals (overt: $\chi^2=22.67$, $p<0.001$, $h=0.685$; NP: $\chi^2=5.30$, $p=0.0213$, $h=0.324$; null: $\chi^2=37.96$, $p<0.001$, $h=0.902$).

These findings support H2. Whereas null pronouns are the tendency in Spanish monolinguals and in upper-advanced learners, the rest of learners differ from monolinguals and show more variability in RE forms. Null pronominal subjects are gradually acquired with proficiency level, whereas overt pronouns show the opposite pattern. Crucially, NPs are a frequent RE form to resolve anaphora in PAS scenarios for both learners and monolinguals. We turn next to the division of labor of such RE forms.

3.3. RQ3/H3: division of labor of the different anaphoric forms

First, we focus on Spanish monolinguals' production to clarify the division of labor in PAS scenarios and to settle the question of whether the alleged flexibility of overt pronouns is more apparent than real. Figure 5 shows a clear bias of null pronouns (93.6%) toward subject antecedents (13), which confirms the PAS and supports most previous research in Spanish. Overt pronouns (32.4%) show a timid bias toward non-subject antecedents, (16b), as previously reported in the literature but, crucially, if we include NPs as a possible RE form, NPs show a strong bias (64.7%) toward non-subject antecedents, (20). Thus, NPs

play an important role in PAS scenarios and this could explain the apparently "flexible" bias found for overt pronouns previously reported.

(20) Él_i acaba rechazándola_j así que la **chica**_j harta de...
[Monolingual: ES_WR_30_3_JVM].

"He_i ends up rejecting her_j, so **the girl**_j, being fed up with..."

Importantly, if we consider overt and NPs forms together (overtly realized REs), then a neater division of labor shows up (Figure 6): null pronouns are biased toward subject antecedents (93.6%) yet overtly realized REs are biased toward non-subject antecedents (97.1%). Thus, corpus data reveals that the division of labor of AR in native Spanish is more complex and more clear-cut than previously assumed since NPs play a key role. These findings explain the division of labor in native Spanish and therefore settle the dispute on the apparent flexibility of overt pronouns in PAS scenarios.

Let us now compare learners against monolinguals regarding the production of RE forms for subject vs. non-subject antecedents. As for subject-antecedent biases (Figure 7), Spanish monolinguals show a clear-cut bias as they produce almost exclusively null pronominal subjects (93.6%). Intermediates show equal variability across all three RE forms (null 35.2%, overt 35.2%, NP 29.6%), as illustrated in (21a, b, c) respectively, and their production is significantly different from monolinguals (null: $\chi^2=44.05$, $p<0.001$, $h=1.358$; overt: $\chi^2=26.09$, $p<0.001$, $h=1.270$; NP: $\chi^2=10.87$, $p=0.0010$, $h=0.638$). From lower advanced to upper advanced we can see an increasing trend toward the native norm, particularly for null pronouns (lower advanced: null 47.8%, NP 26.1%, overt 26.1%; upper advanced: null 75.5%, overt 17.8%, NP 6.7%), though, crucially, each advanced group significantly differs from the monolingual group: lower advanced vs. monolinguals (null: $\chi^2=28.75$, $p<0.001$, $h=1.101$; overt: $\chi^2=18.20$, $p<0.001$, $h=1.072$; NP: $\chi^2=8.07$, $p=0.0045$, $h=0.558$); upper advanced vs. monolinguals (null: $\chi^2=7.00$, $p=0.0081$, $h=0.521$; overt: $\chi^2=11.91$, $p=0.0006$, $h=0.870$; except for NPs, where there are no significant differences $\chi^2=0.00$, $p=0.9646$, $h=0.009$).

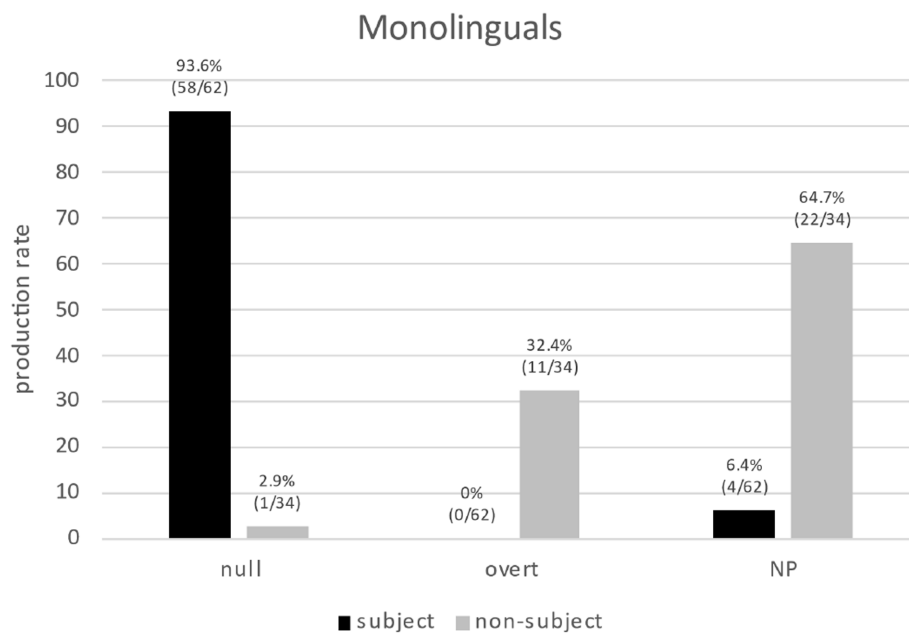


FIGURE 5
Monolinguals' production of REs (null/overt/NP) for subject/non-subject antecedents.

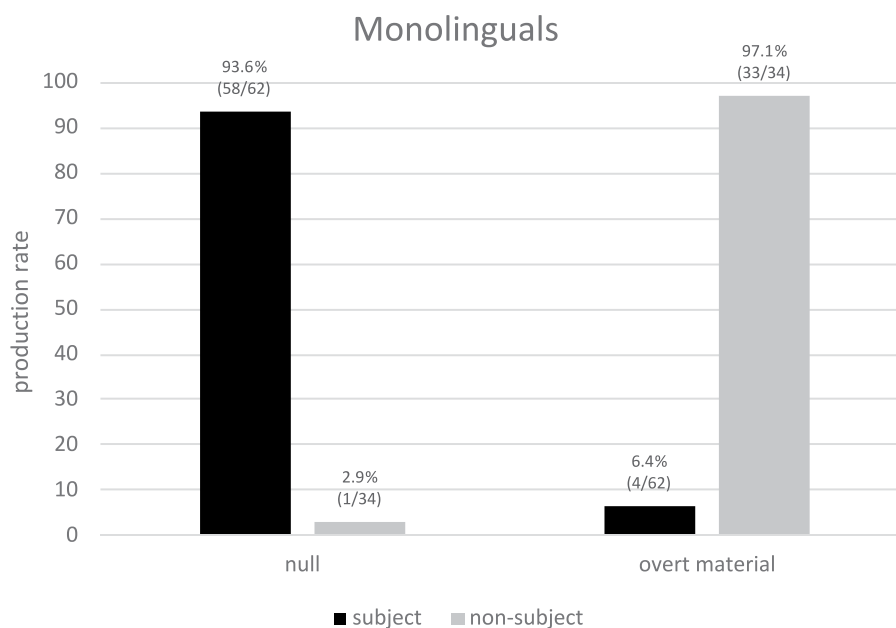


FIGURE 6
Monolinguals' production of REs (null vs. overtly realized REs) for subject/non-subject antecedents.

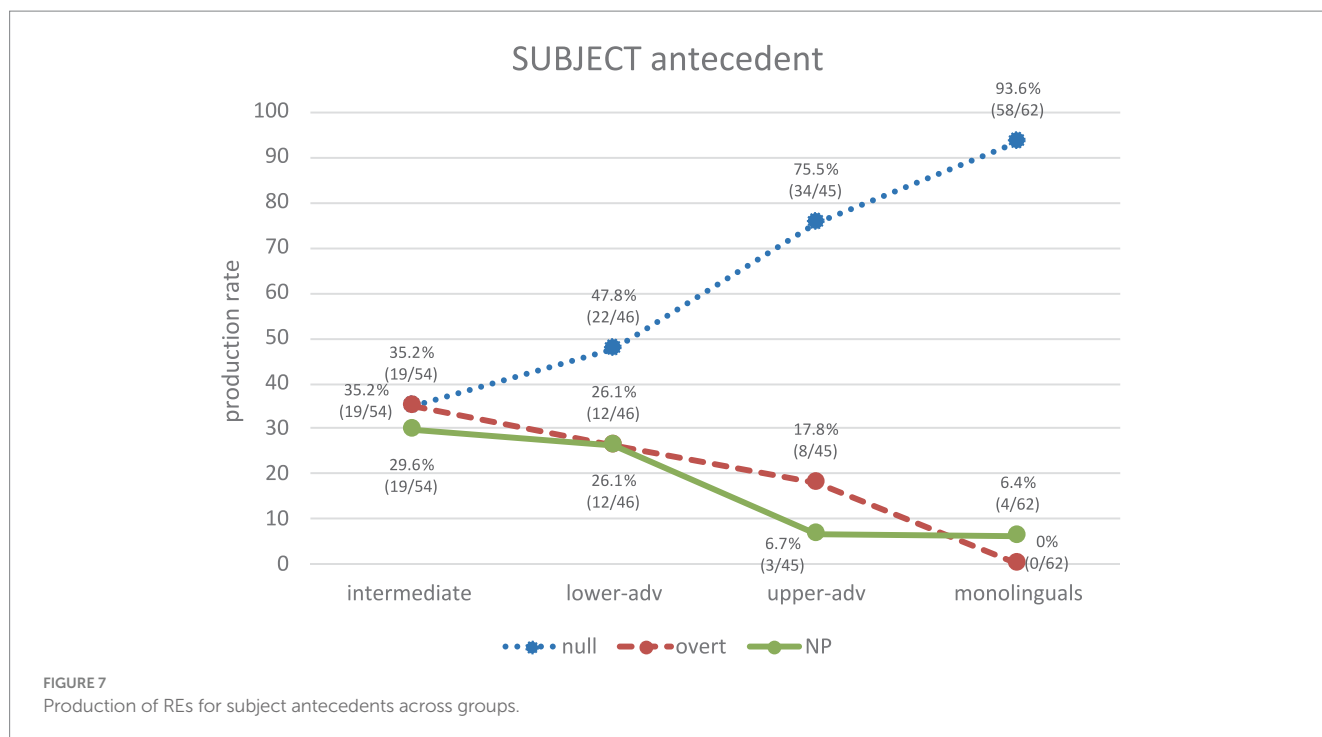
In short, intermediates know that a null pronoun can select a subject antecedent, but they equally produce overt pronouns (as in their L1) and NPs. Nearly half of the productions of lower-advanced learners are null pronouns. Upper-advanced learners show native-like discriminations, but their productions are not fully native-like yet.

(21) a. Brooke_i despide al maestro_j y Ø_i emplea Elle_k... [Learner: EN_WR_31_21_7_3_DNP].

"Brooke_i fires the teacher_j and Ø_i employs Elle_k..."

b. La madre_i es sumisa al padre_j a través de la película. **Ella_i** no ha sabido... [Learner: EN_WR_25_22_17_3_BBB].

"The mother_i is submissive to the father_j throughout the film. **She_i** did not know..."



c. Rose_i quiere a ve Jack_j así que **Rose_i** busca a Jack_j. [Learner: EN_WR_26_18_3_3_BRS].

“Rose_i wants to see Jack_j so **Rose_i** looks for Jack_j.”

Consider now non-subject antecedent biases (Figure 8). Spanish monolinguals' production clearly indicates that NPs (64.7%) (and not overt pronouns) are the privileged RE form to refer to a non-subject antecedent. Crucially, null pronouns are hardly an option for any group, so learners know from the outset that a null pronoun is not an adequate form to refer to a non-subject antecedent. It is therefore remarkable that no null pronouns are used in purely structural PAS configurations. As for learners, overt pronouns and NPs are highly produced, but learners are rather indeterminate about them, particularly intermediates, who show optionality in their production (47.2% overt vs. 52.8% NP), and the two advanced groups, who also show a rather indeterminate pattern where overt pronouns are slightly higher than NPs, as in (22 a, b): lower advanced (56.3% vs. 40.6%), upper advanced (54.6% vs. 40.9%).

(22) a. Bond_i encuentra Vesper_j, y ella_i se disculpa. [Learner: EN_WR_36_19_5_3_MWB].

“Bond_i finds Vesper_j and she_i apologizes.”

b. ...ella_i escribe algunas cartas a Michael_j. Pero Michael_j no responde. [Learner: EN_WR_38_9_30_3_JG].

“...she_i writes some letters to Michael_j. But Michael_j does not reply.”

Figure 8 visually shows that the learners' pattern is either optional (intermediates) or somewhat opposite to the monolinguals' (advanced groups). The low frequencies in production in all groups may explain

why no significant differences are observed between each of the learner groups and the monolinguals ($p > 0.05$ in all cases, though $p < 0.50$ for each of the advanced groups vs. the monolinguals, which represent marginally non-significant differences).

These findings, taken together, support *H3* since null pronouns show a strong bias toward subject antecedents (with learners showing an increasing sensitivity to this), whereas overt material (i.e., overt pronouns as well as NPs) shows a clear bias toward non-subject antecedents.

3.4. RQ4/H4: the syntax-discourse interface

Recall that, at this point, we need to discriminate between purely structural PAS results (RQ 3, previous section) from purely information status/discursive PAS results (RQ4, this section). This will allow us to determine whether the traditional assumption of a correspondence/overlap between syntactic position (subject/non-subject) and information status (topic continuity/shift), as stated in section 1.2, is reflected in production data. Recall that RQ4 will additionally allow us to check for possible deficits at the syntax-discourse interface, as predicted by the IH.

Figure 9 shows the use of REs in topic-continuity contexts, where the production of null pronominal subjects is higher for all groups, although the percentages between groups vary considerably. There is a clear increase of nulls from the intermediate to the monolingual group: intermediate (38.8%), lower-adv (47.8%), upper-adv (76.2%), monolingual (95%). If we compare these results with Figure 7, we can observe a similar trend in the results and a similar statistical behavior. In particular, intermediates show again similar variability across all three RE forms (null 38.8%, overt, 30.6%, NP 30.6%), as shown in (23a-c) respectively and their production is significantly different

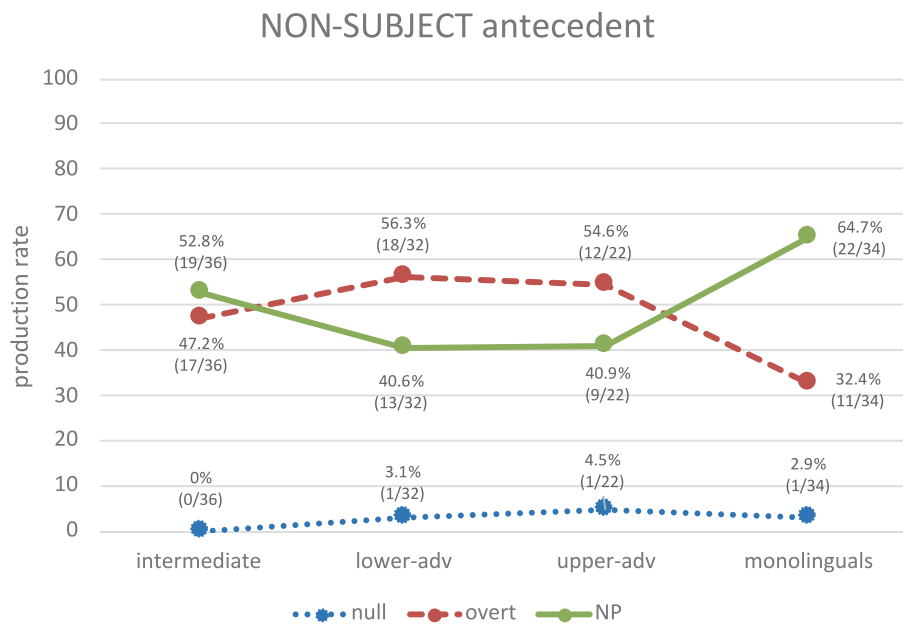


FIGURE 8
Production of REs for non-subject antecedents across groups.

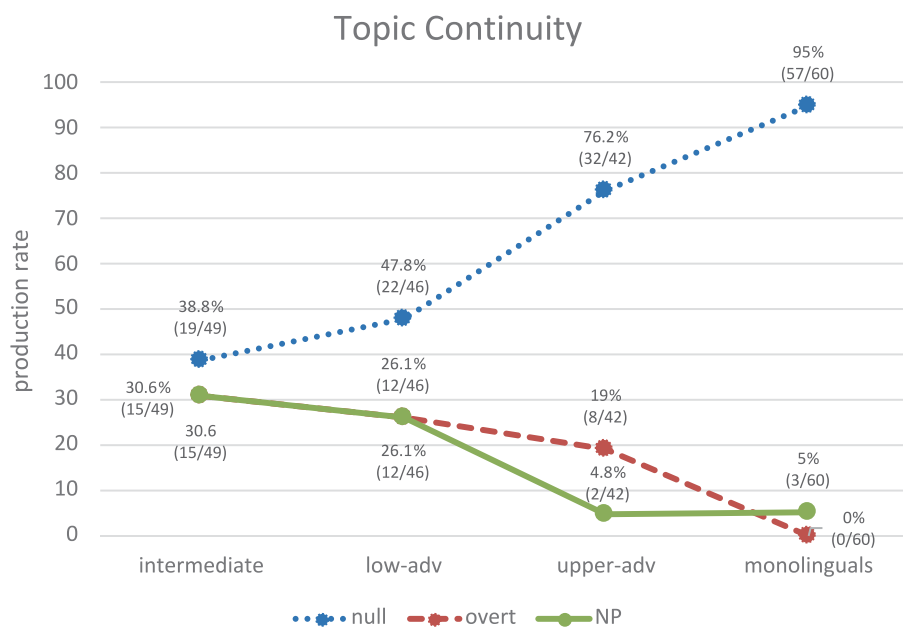


FIGURE 9
Production of REs in topic-continuity contexts across groups.

from monolinguals (null: $\chi^2=40.39$, $p<0.001$, $h=1.346$; overt: $\chi^2=21.30$, $p<0.001$, $h=1.173$; NP: $\chi^2=12.83$, $p=0.0003$, $h=0.772$). From lower advanced to upper advanced we can see again an increase toward the native norm, particularly for null pronouns (lower advanced: null 47.8%, NP 26.1%, overt 26.1%; upper advanced: null 76.2%, overt 19%, NP 4.8%), though, once again, each advanced group significantly differs from the monolingual group: lower advanced vs. monolinguals (null: $\chi^2=30.52$, $p<0.001$, $h=1.163$; overt: $\chi^2=17.65$,

$p<0.001$, $h=1.072$; NP: $\chi^2=9.53$, $p=0.0020$, $h=0.621$); upper advanced vs. monolinguals (null: $\chi^2=7.86$, $p=0.0050$, $h=0.568$; overt: $\chi^2=12.40$, $p=0.0004$, $h=0.903$; except for NPs again, where there are no significant differences ($\chi^2=0.00$, $p=0.9563$, $h=0.011$).

(23) a. Rose_i deja su madre_j y Cal_k y Ø_i va a buscar Jack_i. [Learner: EN_WR_26_18_3_3_BRS].

“Rose_i leaves her mother_j and Cal_k and Ø_i goes to find Jack_i.”

b. Un día el hombre_i estaba sentado en la selva y Ø_i vio la dictadora_j y después **él_i** vio un tigre... [Learner: EN_WR_31_20_Unknown_STS].

“One day the man_i was sitting in the jungle and Ø_i saw the dictatress_j and then he_i saw a tiger...”

c. ... un hombre_i muy rico quiere Satine_j. **El hombre rico_i** tiene mas poder que el hombre pobre_k. [Learner: EN_WR_35_20_10_3_CES].

“A very rich man_i loves Satine_j. **The rich man_i** has more power than the poor man_k.”

By contrast, Figure 10 shows the use of REs in topic-shift contexts. Again, these results show a similar trend to those in Figure 8: monolinguals produce mainly NPs (63.9%), followed by overt pronouns (30.6%). Lower-adv and upper-adv learners show a trend that is rather inverse (though less marked) to monolinguals, by producing overt (58.1, 50%) followed by NPs (41.9, 41.7%), as in (24a, b). Intermediates produce more NPs (54.1%), closely followed by overt (45.9%). Once again, the rather low frequencies in production in all groups may be behind the non-significant differences between each of the learner groups and the monolinguals: non-significant differences ($p > 0.05$) in most contrasts; marginally non-significant differences ($0.05 < p < 0.10$) for NPs in the lower-advanced vs. monolinguals contrast ($\chi^2 = 3.23$) and the upper-advanced vs. monolinguals contrast ($\chi^2 = 2.87$); and only one significant difference for overt pronouns in the lower-advanced vs. monolinguals contrast ($\chi^2 = 5.13$, $p = 0.0234$, $h = 0.561$).

(24) a. ...Ben_i tenía memorias de su esposa_j y su vida con ella_j. Ella_j estaba muy bonita... [Learner: EN_WR_37_18_5_3_JEP].

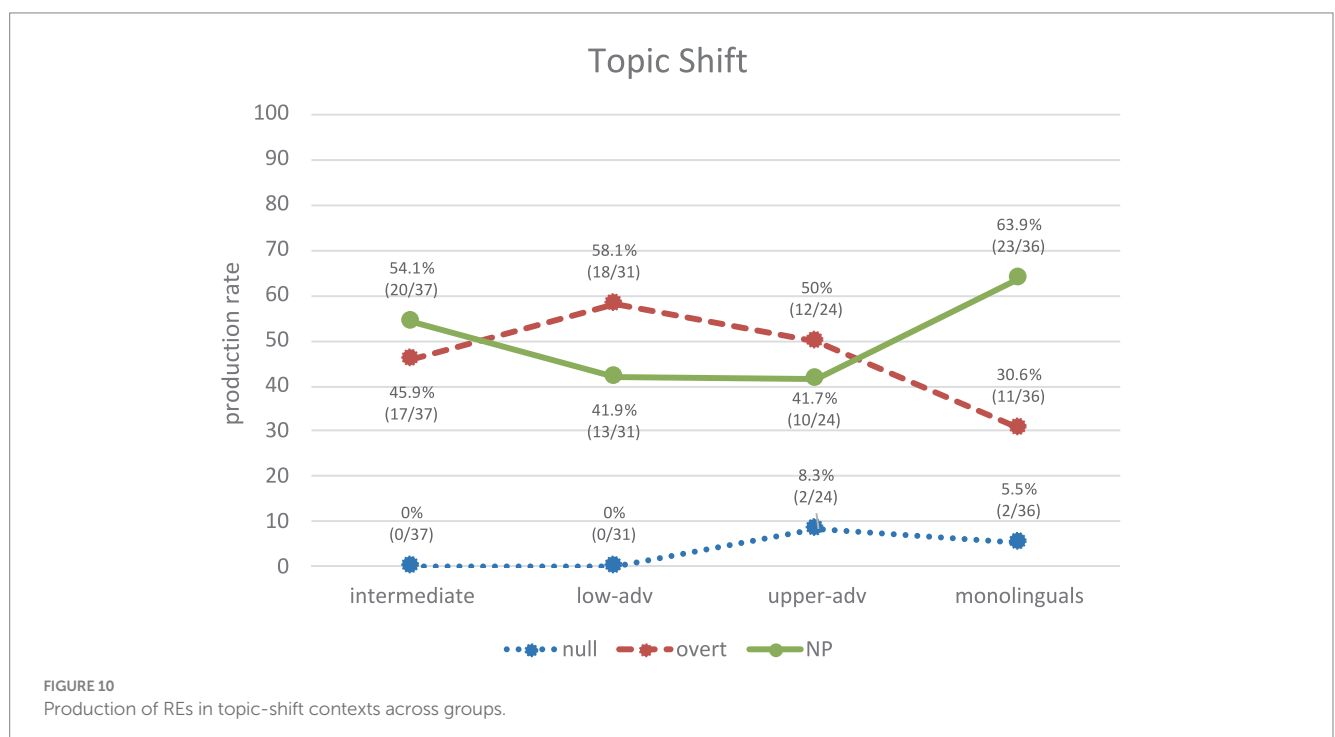
“...Ben_i had memories of his wife_j and his life with her_j. She_j was very pretty...”

b. Pilar_i empieza de desarrollar sus propias opiniones, fuera de su esposo_j. Su esposo_j ha empezado una clase donde Ø_i aprende... [Learner: EN_WR_41_19_5_3_AEM].

“Pilar_i begins to develop her own opinions, outside her husband_j. Her husband_j has started a class where Ø_i learns...”

The results in Figures 9, 10 thus show that syntactic position (subject/non-subject) overlaps with information status (topic continuity/shift) in such a way that null pronouns typically mark a continuation of topic of the subject antecedent, whereas overt material (NPs and overt pronouns) typically marks a shift in topic. These results empirically demonstrate that the traditional experimental assumption in section 1.2 is correct in corpus production data.

When it comes the syntax-discourse interface, recall that Sorace's IH predicts deficits with AR even at advanced levels. This is confirmed in this study, but only partially since our results show that not all syntax-discourse PAS scenarios are equally problematic. In topic-continuity contexts, despite learners' steady increase of null pronouns, the upper-advanced group (76.2%) still significantly differs from monolinguals (95%), but no significant differences were found in topic-shift scenarios with either NPs or overt pronouns. This differential effect is in line with Lozano's (2016) Pragmatic Principles Violation Hypothesis (PPHV), originally proposed for general AR in L1 English-L2 Spanish but also confirmed in other scenarios: AR in L1 Greek-L2 Spanish (Lozano, 2018; Margaza and Gavarró, 2022); AR in L1 English-L2 Spanish and L1 Spanish-L2 English (Quesada, 2021); clitic pronouns in L1 English-L2 Spanish (García-Tejada, 2022); and pragmatic implicatures in L1 Chinese-L2



English (Feng, 2022). The PPVH postulates that learners typically obey the pragmatic Principle of Clarity (i.e., they attain native-like knowledge in topic-shift contexts by using full RE forms to avoid ambiguity) but often violate the Principle of Economy (i.e., they produce overt pronouns in topic continuity, which leads to redundancy).

To summarize, the results showed that the choice of REs depends both on (i) the syntactic position of its antecedent (null→subject vs. NP/overt→non-subject), and (ii) the information status of its antecedent (null→topic continuity vs. NP/overt→topic shift). *H4* is confirmed as there is a correspondence between syntactic position and information structure in PAS. Finally, the PPVH is confirmed since the most advanced L2ers cannot attain full native-like competence in topic-continuity contexts, but they can in topic-shift contexts.

3.5. RQ5/H5: cross-linguistic influence

Recall that a null-subject language like Spanish allows null pronominal subjects in all syntactic configurations (coordination and non-coordination), whereas a non-null subject language like English allows them only in topic continuity *and* coordination. If L1 transfer plays a role in PAS, L1 English-L2 Spanish learners are expected to produce null pronouns mostly in contexts where English allows them.

In topic continuity and coordinate syntactic configurations (cf. (16b)), all groups produce mostly null pronominal subjects. Learners show a slight increasing trend toward the native norm, though only the upper-advanced group shows native-like knowledge (Figure 11): intermediates vs. monolinguals ($\chi^2 = 10.54$, $p = 0.0012$, $h = 1.159$); lower-advanced vs. monolinguals ($\chi^2 = 7.78$, $p = 0.0053$, $h = 0.994$); upper-advanced vs. monolinguals ($\chi^2 = 2.93$, $p = 0.0870$ n.s., $h = 0.613$). By contrast, in topic continuity and non-coordinate configurations

learners' production of null subjects (cf. (17) and (18)) is much lower and is always significantly different from monolinguals': intermediates vs. monolinguals ($\chi^2 = 34.27$, $p < 0.001$, $h = 1.687$); lower-advanced vs. monolinguals ($\chi^2 = 23.97$, $p < 0.001$, $h = 1.463$); upper-advanced vs. monolinguals ($\chi^2 = 5.98$, $p = 0.145$, $h = 0.715$). Additional within-group comparisons¹³ show that Spanish monolinguals' production of null pronouns in topic-continuity coordinate vs. non-coordinate configurations is not significantly different, as expected ($\chi^2 = 3.38$, $p > 0.05$, n.s), whereas learners' production is significantly different: intermediates ($\chi^2 = 16.29$, $p < 0.02$); lower-advanced ($\chi^2 = 13.29$, $p < 0.02$); upper-advanced ($\chi^2 = 5.52$, $p < 0.02$). Results suggest that learners' significantly higher use of null pronouns in coordinate than in non-coordinate configurations reflects L1 English influence. Interestingly, learners show a strong gradual trend toward the native norm (intermediate 15.1%, lower-adv 24%, upper-adv 60%), which suggests their sensitivity to the allowability of null pronouns in non-coordinate scenarios increases with proficiency, though their production rates (even at upper-advanced levels) are far from Spanish monolinguals'. This confirms learners' transfer of null pronouns but an increasing sensitivity to their pragmatics.

3.6. RQ6: sentential configuration

Table 2 shows that that the production of intersentential configurations is two thirds (or higher) the production of

¹³ The latest release of UAM Corpus Tool (version 6.2j, February 2023) does not allow complex within-group comparisons, so we used an earlier release (version 3.3x, August 2021) to do the analysis, though note that version 3.3x reports p value ranges (non significant $p > 0.05$, significant $p < 0.05$, highly significant $p < 0.02$) and does not report effect size.

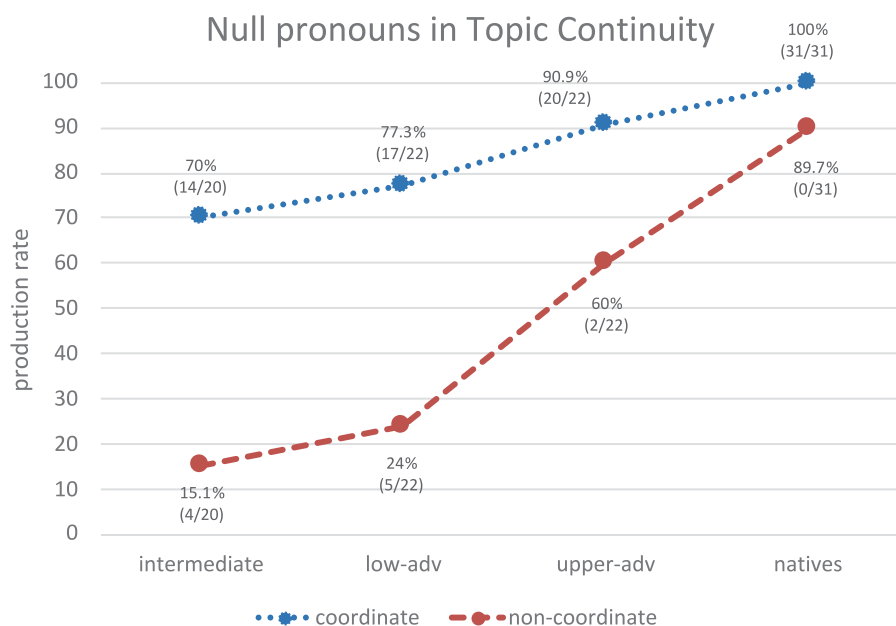
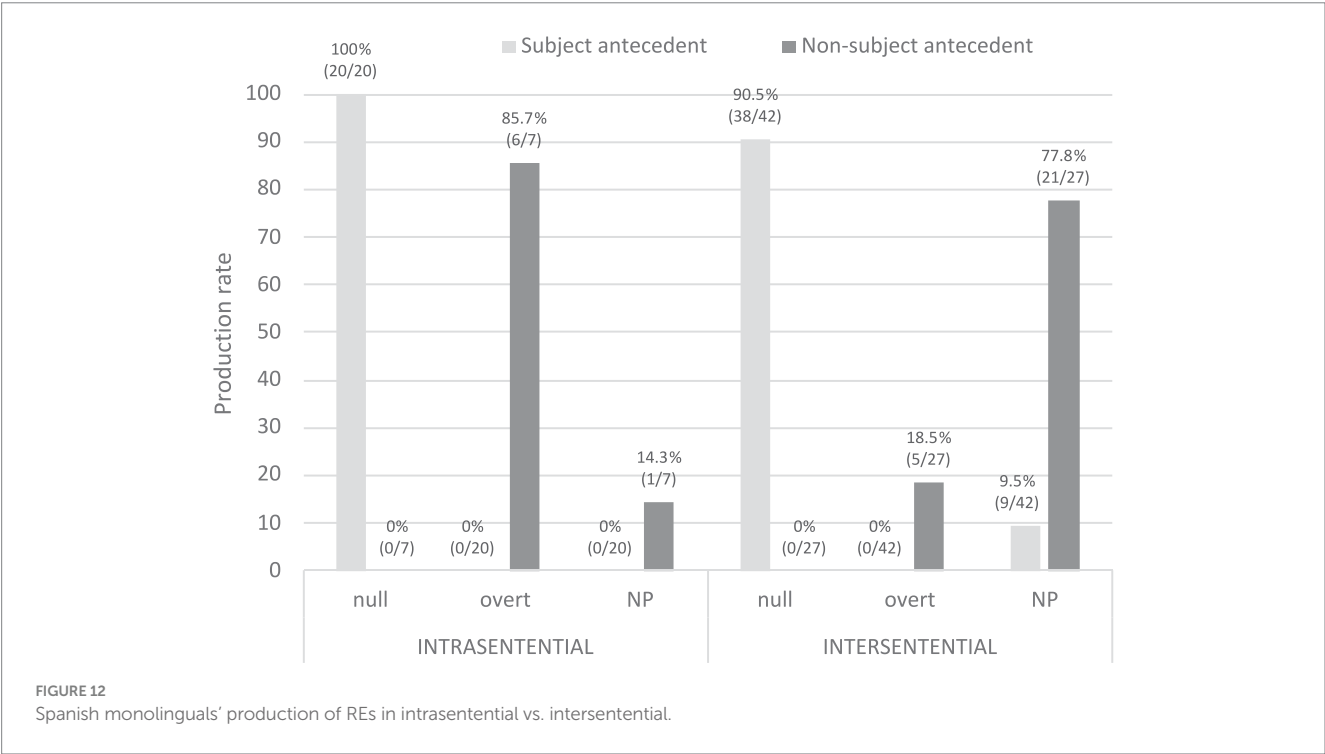


FIGURE 11
Production of REs in topic continuity and coordinate contexts across groups (only null pronouns plotted).

TABLE 2 Syntactic configuration: inter- vs. intra-sentential.

		Intermediate		Lower-adv		Upper-adv		Monolinguals	
INTER-SENT.		79.3% (73/92)		85.7% (72/84)		69.7% (53/76)		64.3% (72/112)	
INTRA-SENT.	Main_subord	20.7% (19/92)	94.74% (18/19)	14.3% (12/84)	83.3% (10/12)	30.3% (23/76)	91.3% (21/23)	35.7% (40/112)	90% (36/40)
	Subord_main		5.26% (1/19)		16.67% (2/12)		8.7% (2/23)		10% (4/40)



intrasentential configurations for both L2ers and monolinguals, which indicates that the most natural sentential configuration for AR in PAS scenarios is intersentential, either independent sentences as in [sentence].[sentence] or coordinate sentences as in [sentence]&[sentence]. This clear-cut trend has been rather overlooked in the design of stimuli in previous experimental studies, where intrasentential configurations like [main [subordinate]] have been typically the focus of attention. Importantly, only the upper-advanced learners can attain native-like competence as they are not significantly different from Spanish monolinguals ($\chi^2=0.60$, $p=0.4371$, $h=0.116$), whereas the intermediates ($\chi^2=5.57$, $p=0.0182$, $h=0.338$), and the lower-advanced learners ($\chi^2=11.31$, $p=0.0008$, $h=0.506$) significantly differ from monolinguals.

Recall that an additional question concerns the order of main and subordinate clauses. Table 2 shows the clausal order for the low-frequency intrasentential configurations: Main-subordinate is overwhelmingly more frequent than subordinate-main for both L2ers and monolinguals. Note that no inferential statistics are performed here due to the low frequencies.

In short, in natural production PAS scenarios are overwhelmingly intersentential and, when they happen to be intra-sentential, the most

frequent clausal order is main-subordinate. This is so in native and non-native grammars. These findings provide clear tips for those researchers wishing to design experimental PAS configurations that intend to look as natural as possible.

A final consideration is whether the sentential configuration is a factor that modulates the choice of RE in PAS in native Spanish (Figure 12). Null pronominal subjects are clearly biased toward subject antecedents regardless of the type of sentence (100% intrasentential, 90.5% intersentential), whereas an interesting subdivision of labor is observed when the bias is toward non-subject antecedents: overt pronouns in intrasentential (85.7%) but NPs in intersentential (77.8%), as shown in (25a, b). In other words, topic continuity (subject bias) is marked via null pronouns irrespective of the sentential configuration, but topic shift (non-subject bias) is marked via overt pronouns intrasententially yet via NPs intersententially, which is a finding not reported in the previous literature. Sentential configuration is therefore an additional factor that modulates the division of labor of REs in PAS configurations in native Spanish. This issue merits further investigation in future studies containing larger frequencies of learner and native corpus data.

- (25) a. Intrasentential: overt pronoun biasing toward a non-subject antecedent:

Marco_i está celoso y Ø_i no se adapta bien a esta nueva vida de Verónica_j, cuando **ella_j** empieza a tomar a sus amigos como amantes. [Learner: EN_WR_42_21_8_3_LBK].

“Marco_i is jealous and Ø_i does not adapt well to this new idea of Veronica_j when **she_j** starts taking her friends as lovers.”

- b. Intersentential: NP pronoun biasing toward a non-subject antecedent:

Ella_i le_j tiene mucho cariño, pero Ø_i se niega a desmentir sus votos para estar con él. **Nacho_j** se deja guiar por un idealismo optimista... [Learner: EN_WR_42_21_10_3_LBK].

“She_i is very fond of him_j, but Ø_i refuses to deny her vows to be with him_j. **Nacho_j** allows himself to be guided by an optimist idealism...”

4. General discussion and conclusion

RQ1 called into question the PAS as a prototypical way of resolving anaphora in native (and L2) Spanish. Our corpus results confirmed a low production of PAS compared to other AR configurations in natural language production. So, as we found during the corpus sample selection (section 2.2), it is difficult to find PAS in natural narrative production and, in those narrations that include PAS, their frequency is rather low. Carminati's (2002) original PAS proposal for Italian has triggered a wealth of experimental studies in many languages and bilingual populations. These studies have blindly tested PAS (and slight variants of it) over and over again but our corpus data show that the PAS is neither a common phenomenon nor prototypical way of resolving anaphora.

Results from RQ2 confirmed the hypothesis that Spanish native discourse contains mainly null pronominal subjects, while learners' production is significantly lower. Importantly, two crucial findings for native Spanish PAS were (i) the rather low production of overt pronouns, which contrasts with their importance in experimental studies, and (ii) the high production of NPs as an anaphoric device, an overlooked factor in experimental studies. Learners' PAS behavior ranged from intermediates' strong influence from their L1 English (overt pronouns and NPs predominate, with low rates of null pronouns), the indeterminacy of lower advanced learners (production of one third of each RE form), and difficulty to attain native levels by the upper-advanced group since they still produce significantly more overt pronouns than monolinguals, in line with previous corpus research on L2 Spanish dealing with AR in general (Montrul and Rodríguez Louro, 2006; Lozano, 2009, 2016). These findings become more meaningful when we incorporate syntax-discourse factors in PAS, as we will discuss below.

RQ3 addressed a much-debated topic in the literature on native Spanish: the division of labor of RE forms in PAS. Experimental studies report a clear role for null pronouns (they show a strong subject-antecedent bias), yet overt pronouns show a “flexible” behavior (non-subject- as well as subject-antecedent biases). The corpus data showed a clear division of labor when we consider overtly realized REs together (i.e., overt pronouns and NPs): null

pronouns clearly select subject antecedents whereas overt REs clearly select non-subject antecedents. This is quite revealing as NPs were not typically considered in previous experimental PAS studies (except for Gelormini-Lezama and Almor, 2011). The relevance of corpus data then becomes clear as a complementary (and needed) source of evidence for experimental data in the study of bilingualism.

As for learners' subject antecedents, they start off by showing indeterminacy and no clear PAS strategy in L2 Spanish, but then show a gradual development toward the native norm, but even the upper-advanced group still significantly produces more overt pronouns (and less null pronouns) than monolinguals do to refer to the subject. The results are in line with previous studies regarding development (Jegerski et al., 2011) and native-like knowledge but lack of full native-like attainment at advanced levels (Bel et al., 2016b; Clements and Domínguez, 2017). As for learners' non-subject antecedents, if we consider overt pronouns and NPs together, the bias is clearer for all groups as overt REs are biased toward non-subject antecedents. So, it seems that the division of labor in learners' is clearer from early stages for non-subject antecedents than for subject antecedents. This is not surprising as the antecedent bias is somehow related to the information status (i.e., topic continuity/shift) and topic continuity is more problematic than topic shift, as we discuss next.

Results for RQ4 confirmed the correspondence between information status and syntactic configuration (i.e., null pronouns → subject antecedent/topic continuity; overt pronouns & NPs → non-subject antecedent/topic shift). Regarding the deficits at the syntax-discourse interface predicted by the IH (Sorace, 2011), learners showed deficits, but there were differential effects, as predicted by the PPVH (Lozano, 2016): Learners showed native-like behavior in topic-shift, but not in topic-continuity contexts, where even upper-advanced learners redundantly use overt pronouns. In short, learners are more redundant than ambiguous with the PAS.

As for RQ5, learners' lack of native-like attainment with PAS is also motivated by transfer of null pronominal subjects from their L1 in topic continuity and coordination (and not in topic continuity and non-coordination), a fact also reported by Martín-Villena and Lozano (2020) for diverse AR contexts. Curiously, the cross-linguistic effect is milder in the opposite direction (L1 Spanish-L2 English), as reported by Quesada and Lozano (2020), so future research could investigate this asymmetry in a more controlled way, e.g., by keeping the task and the type of AR analysis constant but turning the language pairs (L1 English-L2 Spanish vs. L1 Spanish-L2 English) into a variable. Despite transfer, our results also show acquisition effects since learners gradually increase their production of null pronouns in both contexts as their proficiency increases.

As for RQ6, our corpus data showed that 1/3 of PAS configurations were intrasentential, of which over 90% were main-subordinate order. Interestingly, some of the studies reviewed above that investigated intrasentential sentences showed contradictory results depending on the order of presentation: main-subordinate order (Chamorro et al., 2016; Bel et al., 2016a) vs. subordinate-main order (Filiaci, 2010; Filiaci et al., 2014; Keating et al., 2016) (cf. online Supplementary material for exact details). Importantly, our corpus findings also show that null pronouns are clearly biased toward subject antecedents regardless of the type of sentential configuration, but for non-subject antecedents the configuration modulates the choice of RE: overt

pronouns are biased toward non-subject antecedents intrasententially whereas NPs do so in intersententially.

The current study presents certain limitations. A larger corpus sample would have probably yielded more stable findings but recall our difficulty in finding texts containing enough PAS examples. Additionally, the tasks certainly lead speakers to narrate different films or describe different famous people, which generates a wide and heterogeneous variety of AR scenarios in the texts produced. This could be minimized by using a prompted task (e.g., the narration of a short Charles Chaplin video clip).

Our findings show the relevance of learner corpus research to investigate theoretically-motivated L2 phenomena (Lozano, 2021b). Corpus data have uncovered certain key factors that could be certainly implemented in future experiments. Our research group is currently implementing some of these factors into new experiments (NPs as a form of RE, number of potential antecedents, antecedent-anaphor distance, etc). This is in line with recent claims (Mendikoetxea and Lozano, 2018; Gilquin, 2021) that the triangulation of *experimental* and *corpus* methods leads to a more well-rounded understanding of complex linguistic phenomena in bilingualism and SLA.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found at: <http://cedel2.learnercorpora.com>.

Ethics statement

The studies involving humans were approved by Comité en investigación Humana (Universidad de Granada), 1794/CEIH/2020. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

References

- Alonso-Ovalle, L., Fernández-Solera, S., Frazier, L., and Clifton, C. (2002). Null vs. overt pronouns and the topic-focus articulation in Spanish: 2704. *Ital. J. Linguist.* 14, 151–170.
- Bel, A., and García-Alcaraz, E. (2015). "Subject pronouns in the L2 Spanish of Moroccan Arabic speakers: evidence from bilingual and second language learners" in *The Acquisition of Spanish in understudied language pairings*. eds. T. Judy and S. Perpiñán (Amsterdam, John Benjamins).
- Bel, A., García-Alcaraz, E., and Rosado, E. (2016a). Reference comprehension and production in bilingual Spanish: the view from null subject languages. In FuenteA. A. de la, E. Valenzuela and C. Martínez Sanz (Eds.), *Language acquisition beyond parameters* Amsterdam: John Benjamins.
- Bel, A., Sagarra, N., Comínguez, J. P., and García-Alcaraz, E. (2016b). Transfer and proficiency effects in L2 processing of subject anaphora. *Lingua* 184, 134–159. doi: 10.1016/j.lingua.2016.07.001
- Carminati, M. N. (2002). *The processing of Italian subject pronouns* Boston: University of Massachusetts.
- Chamorro, G., Sorace, A., and Sturt, P. (2016). What is the source of L1 attrition? The effect of recent L1 re-exposure on Spanish speakers under L1 attrition. *Biling. Lang. Cogn.* 19, 520–532. doi: 10.1017/S1366728915000152
- Clahsen, H., and Felser, C. (2006). How native-like is non-native language processing? *Trends Cogn. Sci.* 10, 564–570. doi: 10.1016/j.tics.2006.10.002
- Clements, M., and Domínguez, L. (2017). Reexamining the acquisition of null subject pronouns in a second language: focus on referential and pragmatic constraints. *Linguist. Approach. Biling.* 7, 33–62. doi: 10.1075/lab.14012.cle
- Egbert, J., Larsson, T., and Biber, D. (2020). *Doing linguistics with a corpus: Methodological considerations for the everyday user*. Cambridge: Cambridge University Press.
- Feng, S. (2022). L2 tolerance of pragmatic violations of informativeness: evidence from ad hoc implicatures and contrastive inference. *Linguist. Appr. Bil.* doi: 10.1075/lab.21064.fen
- Fernández, E. M., and Smith Cairns, H. (2011). *Fundamentals of psycholinguistics*. Hoboken: Wiley-Blackwell.
- Filiaci, F. (2010). "Null and overt subject biases in Spanish and Italian: a cross-linguistic comparison" in *Selected proceedings of the 12th Hispanic linguistics symposium*. eds. C. Borgonovo, M. Español-Echevarría and P. Prévost (Hoboken: Cascadia Press)
- Filiaci, F., Sorace, A., and Carreiras, M. (2014). Anaphoric biases of null and overt subjects in Italian and Spanish: a cross-linguistic comparison. *Lang. Cogn. Neurosci.* 29, 825–843. doi: 10.1080/01690965.2013.801502
- Flores-Ferrán, N. (2004). Spanish subject personal pronoun use in New York City Puerto Ricans: can we rest the case of English contact? *Lang. Var. Chang.* 16, 49–73. doi: 10.1017/S0954394504161048
- García-Alcaraz, E., and Bel, A. (2019). Does empirical data from bilingual and native Spanish corpora meet linguistic theory? The role of discourse context in variation of subject expression. *Appl. Linguist. Rev.* 10, 491–515. doi: 10.1515/applrev-2017-0101
- García-Tejada, A. (2022). Direct object anaphora resolution in L1 English-L2 Spanish: referring clitics and DPs. *Revista Española de Lingüística Aplicada*. doi: 10.1075/resla.22014.gar

Author contributions

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

Funding

This study was funded by grant number PID2020-113818GB-I00 from MCIN (Ministerio de Ciencia e Innovación), AEI (Agencia Estatal de Investigación) (DOI: <https://doi.org/10.13039/501100011033>) to CL. Publication fees were paid by MCIN/AEI and by the Department of English and German Philology (Universidad de Granada).

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1246710/full#supplementary-material>

- Geber, D. (2006). Processing subject pronouns in relation to non-canonical (quirky) constructions. *Ottawa Pap. Linguist.* 34, 47–61.
- Gelomini-Lezama, C., and Almor, A. (2011). Repeated names, overt pronouns, and null pronouns in Spanish. *Lang. Cogn. Process.* 26, 437–454. doi: 10.1080/01690965.2010.495234
- Gilquin, G. (2021). “Combining learner corpora and experimental methods” in *The Routledge handbook of second language acquisition and corpora*. eds. N. Tracy-Ventura and M. Paquot (London: Routledge)
- Jegerski, J., Van Patten, B., and Keating, G. D. (2011). Cross-linguistic variation and the acquisition of pronominal reference in L2 Spanish. *Second. Lang. Res.* 27, 481–507. doi: 10.1177/0267658311406033
- Keating, G. D., Jegerski, J., and Vanpatten, B. (2016). Online processing of subject pronouns in monolingual and heritage bilingual speakers of Mexican Spanish. *Biling. Lang. Cogn.* 19, 36–49. doi: 10.1017/S1366728914000418
- Keating, G. D., Vanpatten, B., and Jegerski, J. (2011). Who was walking on the beach?: anaphora resolution in Spanish heritage speakers and adult second language learners. *Stud. Second. Lang. Acquis.* 33, 193–221. doi: 10.1017/S0272263110000732
- Kraš, T. (2008a). Anaphora resolution in Croatian: psycholinguistic evidence from native speakers. In M. Tadić, M. Dimitrova-Vulchanova and S. Koeva (Eds.), *Proceedings of the Sixth International Conference “Formal Approaches to South Slavic and Balkan languages Zagreb, Croatia: Croatian Language Technologies Society – Faculty of Humanities and Social Sciences*.
- Kraš, T. (2008b). “Anaphora resolution in near-native Italian grammars: evidence from anaphora resolution in Spanish heritage speakers and adult second language learners. David (Hoboken: John Benjamins)
- Lozano, C. (2009). “Selective deficits at the syntax-discourse interface: evidence from the CEDEL2 corpus” in *Representational deficits in SLA: Studies in honor of Roger Hawkins*. eds. N. Snape, Y. I. Leung and M. Sharwood Smith (Hoboken: John Benjamins)
- Lozano, C. (2016). “Pragmatic principles in anaphora resolution at the syntax-discourse interface: advanced English learners of Spanish in the CEDEL2 corpus” in *Spanish learner corpus research: State of the art and perspectives*. ed. M. Alonso-Ramos (Hoboken: John Benjamins)
- Lozano, C. (2018). The development of anaphora resolution at the syntax-discourse interface: pronominal subjects in Greek learners of Spanish. *J. Psycholinguist. Res.* 47, 411–430. doi: 10.1007/s10936-017-9541-8
- Lozano, C. (2021a). “Anaphora resolution in second language acquisition” in *Oxford bibliographies in linguistics*. ed. M. Aronoff (Oxford: Oxford University Press)
- Lozano, C. (2021b). “Generative approaches” in *The Routledge handbook of second language acquisition and corpora*. eds. N. Tracy-Ventura and M. Paquot (London: Routledge)
- Lozano, C. (2022). CEDEL2: design, compilation and web interface of an online corpus for L2 Spanish acquisition research. *Second. Lang. Res.* 38, 965–983. doi: 10.1177/02676583211050522
- Mac Donald, M. (2013). How language production shapes language form and comprehension. *Front. Psychol.* 4:226. doi: 10.3389/fpsyg.2013.00226
- Margaza, P., and Gavarró, A. (2022). The distribution of subjects in L2 Spanish by Greek learners. *Front. Psychol.* 12:794587. doi: 10.3389/fpsyg.2021.794587
- Martin-Villena, F. (2023). L1 morphosyntactic attrition at the early stages: Evidence from production, interpretation, and processing of subject referring expressions in L1 Spanish-L2 English instructed and immersed bilinguals [PhD dissertation, Universidad de Granada]. Available at: <https://hdl.handle.net/10481/81920>
- Martin-Villena, F., and Lozano, C. (2020). “Anaphora resolution in topic continuity: evidence from L1 English-L2 Spanish data in the CEDEL2 corpus” in *Referring in a second language: Studies on reference to person in a multilingual world*. eds. J. Ryan and P. Crosthwaite (London: Routledge)
- Mendikoetxea, A., and Lozano, C. (2018). From corpora to experiments: methodological triangulation in the study of word order at the interfaces in adult late bilinguals (L2 learners). *J. Psycholinguist. Res.* 47, 871–898. doi: 10.1007/s10936-018-9560-0
- Montrol, S., and Rodríguez Louro, C. (2006). “Beyond the syntax of the null subject parameter: a look at the discourse-pragmatic distribution of null and overt subjects by L2 learners of Spanish” in *Language acquisition and language disorders*. eds. V. Torrens and L. Escobar (Hoboken: John Benjamins)
- O'Donnell, M. (2009). “The UAM corpus tool: software for corpus annotation and exploration” in *Applied linguistics now: Understanding language and mind/La Lingüística Aplicada actual: Comprendiendo el Lenguaje y la Mente*. eds. C. M. Bretonet al. (Almería: Universidad de Almería)
- Papadopoulou, D., Peristeri, E., Plemenou, E., Marinis, T., and Tsimpli, I. (2015). Pronoun ambiguity resolution in Greek: evidence from monolingual adults and children. *Lingua* 155, 98–120. doi: 10.1016/j.lingua.2014.09.006
- Pickering, M. J., and van Gompel, P. G. (2006). “Syntactic parsing” in *Handbook of psycholinguistics*. eds. M. Traxler and M. A. Gernsbacher. 2nd ed (Cambridge: Academic Press)
- Prentza, A., and Tsimpli, I.-M. (2013). Resolution of pronominal ambiguity in Greek: syntax and pragmatics. *Stud. Greek Linguist.* 33, 197–208.
- Quesada, T. (2021). Studies on anaphora resolution in L1 Spanish-L2 English and L1 English-L2 Spanish adult learners: Combining corpus and experimental methods [PhD dissertation, Universidad de Granada]. Available at: <http://hdl.handle.net/10481/72052>
- Quesada, T., and Lozano, C. (2020). Which factors determine the choice of referential expressions in L2 English discourse? A multifactorial study from the COREFL corpus. *Stud. Second. Lang. Acquis.* 42, 959–986. doi: 10.1017/S0272263120000224
- Shin, N. L., and Smith Cairns, H. (2012). The development of NP selection in school-age children: reference and Spanish subject pronouns. *Lang. Acquis.* 19, 3–38. doi: 10.1080/10489223.2012.633846
- Sorace, A. (2011). Pinning down the concept of “interface” in bilingualism. *Linguist. Approach. Bilingual.* 1, 1–33. doi: 10.1075/lab.1.1.01sor
- University of Wisconsin. (1998). *The University of Wisconsin College-Level Placement Test: Spanish (grammar) form 96M*. Madison: University of Wisconsin Press.



OPEN ACCESS

EDITED BY

Bradley Hoot,
DePaul University, United States

REVIEWED BY

Olga Kellert,
University of A Coruña, Spain
Lindsay Butler,
Boston University, United States

*CORRESPONDENCE

Melanie Uth

✉ uth@uni-potsdam.de

RECEIVED 31 August 2023

ACCEPTED 27 December 2023

PUBLISHED 22 February 2024

CITATION

Uth M, Gutiérrez-Bravo R and Fließbach J
(2024) On the incompatibility of object
fronting and progressive aspect in Yucatec
Maya. *Front. Lang. Sci.* 2:1286520.
doi: 10.3389/flang.2023.1286520

COPYRIGHT

© 2024 Uth, Gutiérrez-Bravo and Fließbach.
This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that
the original publication in this journal is cited,
in accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

On the incompatibility of object fronting and progressive aspect in Yucatec Maya

Melanie Uth^{1*}, Rodrigo Gutiérrez-Bravo² and Jan Fließbach¹¹Institut für Romanistik, University of Potsdam, Brandenburg, Germany, ²Centro de Estudios Lingüísticos y Literarios, El Colegio de México, Mexico City, Mexico

In this paper, we present data from an elicitation study and a corpus study that support the observation that the Yucatec Maya progressive aspect auxiliary *táan* is replaced by the habitual auxiliary *k* in sentences with contrastively focused fronted objects. Focus has been extensively studied in Yucatec, yet the incompatibility of object fronting and progressive aspect in Yucatec Maya remains understudied. Both our experimental results and our corpus study point in the direction that this incompatibility may very well be categorical. Theoretically, we take a progressive reading to be derived from an imperfectivity operator in combination with a singular operator, and we propose that this singular operator implicates the negation of event plurality, leading to an exhaustive interpretation which ranks below corrective focus on a contrastive focus scale. This means that, in a sentence with object focus fronting, the use of the marked auxiliary *táan* (as opposed to the more general *k*) would trigger two contrastive foci, which would be an unlikely and probably dispreferred speech act.

KEYWORDS

fronting, progressive, Yucatec Maya, habitual, focus

1 Introduction

In this paper, we present and account for data from an elicitation study and a corpus study that support the observation that the Yucatec Maya progressive aspect morpheme *táan* is replaced by the imperfective morpheme *k* in sentences with fronted objects (Bohnenmeyer, 2002, p. 265). We suggest that this restriction is the result of a focal component in the meaning of the progressive marker that leads to the exclusion of event plurality. Yucatec Maya (YM), a head-marking Mayan language mainly spoken in south-east Mexico and in parts of Belize (iso code: yua), is traditionally considered a VOS language (but see Gutiérrez-Bravo and Monforte, 2010 for a different view) with robust constituent fronting to the left of the finite verb for the purpose of focusing (1a.–c.).¹

- (1) a. [FOC *Leti*] *kíin-s-ej-ø*.
3.SG die.AF-CAUS-PRF-ABS.3SG
“HE killed him.” (MDG-B: 26)
(Gutiérrez-Bravo, 2017, p. 6)

¹All Yucatec examples are presented according to the 1984 orthographic conventions of the *Comisión de Difusión del Alfabeto Maya* of the *Academia de la Lengua Maya de Yucatán* (CDAM, 1984) and do not necessarily reflect their phonetic form accurately. In this orthographical system, symbols have their expected values except for *ch*=[tʃ], *j*=[h], *x*=[ʃ], and *'*=[ʔ]. The abbreviations used in the examples are the following: ABS: absolutive, AF: agent focus, CAUS: causative, CL: clitic, CP: completive, EP: epenthesis, ERG: ergative, FOC: focus (adverb), HAB: habitual, IND: indicative, NEG: negation, NFP: non-finite passive, PART: participle, PASS: passive, PL: plural, PREP: preposition, PRF: perfect, PRG: progressive, SG: singular, TOP: topic, TR: transitive, TRM: terminative. The glosses ERG and ABS correspond to what is known in traditional Mayan linguistics as the *A* and *B* pronominal series.

- b. *Tumen to'on=e'* [FOC *maaya*] *k*
 because 1.PL-TOP Maya HAB.ERG.1.PL
t'an-ik-ø.
 speak-IND-ABS.3SG
 "Because we, we speak MAYA." (MTK: 63)
 (Gutiérrez-Bravo, 2017, p. 12)
- c. *Pero ma'* [FOC *ti'* *tuláaka*] *k=u*
 but NEG PREP all HAB=ERG.3
ts'a'ab-al=i'.
 give.PASS-IND-CL
 "But it is not granted TO EVERYONE." (MDG-B: 62)
 (Gutiérrez-Bravo, 2015a, p. 21)

Additionally, Yucatec has a particular verb focus construction. In this construction a nonfinite form of the verb occupies the same left peripheral position as other kinds of foci. Furthermore, a dummy verb *beet* "do" functions as the finite (main) verb of the clause, a syntactic process that is not observed elsewhere in the syntax of Yucatec Maya (2.a., b.).

- (2) a. *Okol-bil u beet-ik-ø*
 steal-NFP ERG.3 do-IND-ABS.3SG
wal=e'.
 perhaps=CL
 "Maybe he used to STEAL it."
 (Gutiérrez-Bravo, 2015a, p. 23)
- b. *Bin k=u beet-ik-ø*.
 go HAB=ERG.3 do-IND-ABS.3.SG
 "What he used to do was to GO (away)."
 (Gutiérrez-Bravo, 2015a, p. 25)

Thus, like other Mayan languages, Yucatec consistently shows movement of a focused constituent to the left edge of the clause. However, based on data from a recent elicitation study of object focus constructions with 10 Yucatec speakers from Quintana Roo, Mexico, we will show in what follows that there seems to exist a restriction on focus structures in Yucatec Maya, such that focus fronting is incompatible with the progressive aspect, which is marked with the auxiliary *táan*. Some examples of this auxiliary are presented in (3).²

- (3) a. *Táan k meyaj...*
 PRG ERG.1SG work
 "We were working." (MTK: 106)
- b. *Táan in w-il-ik-ø*.
 PRG ERG.1SG EP-SEE-IND-ABS.3.SG
 "I'm seeing it." (MTK: 85)

Yucatec Maya has a considerable number of aspect and mood auxiliaries and there is nothing in the morpho-syntactic or syntactic properties of *táan* that make it different from any of the other auxiliaries in this language (see Bricker et al., 1998; Bohnemeyer, 2002, for instance). As can be seen in the examples in (3) and (4a.), for example, it has the same distribution as any other auxiliary, i.e., immediately to the left of the ergative

clitic that cross-references the subject. Our results, however, show that it is very different from other auxiliaries with respect to focus fronting.

Focus fronting has been extensively studied in Yucatec (to the point where it is probably the syntactic phenomenon most studied in this language: Tonhauser, 2003; Lehmann, 2008; Gutiérrez-Bravo and Monforte, 2011; Skopeteas and Verhoeven, 2012; Gutiérrez-Bravo, 2015a,b; Verhoeven and Skopeteas, 2015, *inter alia*), yet this particular restriction has gone mostly unnoticed. One exception is Bohnemeyer (2002, p. 265), where it is observed that the habitual auxiliary (referred to as imperfective by this author and others) tends to replace the progressive in all focus constructions. As we show in detail in what follows, our data corroborates this statement for fronted objects to the point that we did not find a single example of *táan* co-occurring with fronted objects, neither in our elicitation study nor in our corpus study (see 4a. vs. b.),³ whereas we found a fairly high number of object fronting with the bound habitual auxiliary *k* and other aspect and mood auxiliaries in our data (see 5a.–c. and Section 3.2).

- (4) a. *Tun man-ik-ø kib (le*
 PRG.ERG.3 buy-IND-ABS.3.SG candle DET
ko'olel=o')
 woman=CL
 "The woman is buying candles."
- b. **kib tun man-ik-ø*
 candle PRG.ERG.3 buy-IND-ABS.3.SG
 (le *ko'olel=o')*
 DET woman=CL
 (Target meaning: "The woman is buying CANDLES.")
- (5) a. *kib k=u*
 candle HAB=ERG.3
man-ik-ø le
 buy-IND-ABS.3.SG DET
ko'olel=o' (ASP = habitual)
 woman=CL
 "The woman buys CANDLES."
- b. *siete in paalal ts'o'ok in*
 seven my children TRM ERG.1.SG
nuukkíin-s-ik-ø. (ASP = terminative)
 bring.up-CAUS-IND-ABS.3.SG
 "I have already brought up SEVEN CHILDREN (of mine)" (MDG-B: 32)

2 This auxiliary has a number of portmanteau allomorphs which incorporate the ergative pronominal clitics, i.e. *túun/tun* for third person [as in (4)], *tíin/tin* for first person singular, etc. See for instance Bohnemeyer (2002: 103).

3 The asterisk (*) in (4b) is, so far, supposed to reflect the non-existence of the PROG marker with fronted objects in our data. Since our studies did not produce negative evidence, future experiments will need to test whether this restriction is as categorical as suggested by our data. We thank an anonymous reviewer for stressing this point. As will be discussed in more detail in Section 4, note that, although limited in their scope, our clear-cut results do still point to the possibility that the incompatibility between focus fronting and progressive aspect may very well be categorical, whereas there is no hint whatsoever in our data that would motivate us to pursue the opposite hypothesis.

- c. *Chéen u tsímin t=u*
 only his horse CP=ERG.3
bi-s-aj-ø. (ASP = completive)
 go-CAUS-PRF-ABS.3SG
 “He brought along only HIS HORSE.” (MTK: 85)

As will be detailed in Section 4, we develop an analysis of these facts that argues for a focal meaning component for *táan* that interacts with the focal meaning of object fronting to render their combination unlikely and probably unacceptable in most, if not all contexts. Our account is based on the semantics of habitual and progressive according to Ferreira (2016) and the idea of scalar implicature as a focus related effect according to Rooth (1992). The outline of the rest of paper is as follows. In Section 2, we provide a brief description of focus in Yucatec, with the aim to establish the necessary conditions to identify syntactic focus constructions in this language. Afterwards, we present our two data sets and the main results of the corresponding word order analyses (Section 3). We then elaborate our semantic framework to account for the above-mentioned effects in terms of mereological semantics and scalar implicatures (Section 4), and we discuss the implications and possible short-comings of our account (Section 5). Finally, Section 6 summarizes our main findings and conclusions.

2 Focus in Yucatec Maya

In Yucatec Maya, contrastively⁴ focused constituents are placed immediately to the left of the auxiliary or verb, as in many other Mayan languages. In (6) and (7), a contrastively focused intransitive subject is fronted.

- (6) [FOC *Teech*] *a w-ojel.*
 2.SG ERG.2 EP-know
 “YOU will know.” (MDG-B: 136)
- (7) [FOC *Leti*] *k=u y-awat.*
 3.SG HAB=ERG.3 EP-scream
 “HE is the one who screams.” (MDG-B: 65)

In example (8), the direct object is focused. As mentioned above, this also brings with it object fronting, since the unmarked order between verb and object in Yucatec is VO. This example furthermore shows an important property of the preverbal field in Yucatec, namely, that this language has a specific topic position to the left of the position occupied by the focused constituent. As in many other Mayan languages, these two positions show different syntactic and prosodic properties, with a high tonal target at the edge of the focused constituent as one cue for the prosodic break between the topic and the focus position (see Verhoeven and Skopeteas, 2015).

- (8) *Tumen to'on=e'* [FOC *maaya*] *k*
 because 1.PL=TOP maya HAB.ERG.1.PL
t'an-ik-ø.
 speak-IND-ABS.3SG
 “Because we, we speak MAYA.” (GRISELDA, p. 1)

The focused constituent can be preceded immediately by a negation, as in example (9). In contrast, topics always appear to the left of the negation. This means that the difference between the relative ordering of topics and foci with respect to negation is also a diagnostic and can be used to distinguish them.

- (9) *To'on=e', ma'* [FOC *leti*]
 1PL=TOP NEG 3SG
kaan-s-a'an-ø to'on=i'.
 learn-CAUS-PART-ABS.3SG 1PL=CL
 “It was not THAT that was taught to us.” (MDG-B:266)

Many Mayan languages show a verb form that is different from the canonical verb form when the transitive subject is focused. In the literature on Mayan languages, this distinct verbal form is known as the agent focus form of the verb. Yucatec is one of the Mayan languages that show this phenomenon. Agent focus constructions are different from most other canonical transitive matrix constructions in the language in that they do not show any kind of auxiliary. The ergative proclitic that agrees with the subject/agent is also absent, and so the subject/agent instead has to be realized as a pronoun or a full noun phrase (Bricker, 1979; Tonhauser, 2003). The mood morphology of the agent focus construction is also different from that observed in canonical transitive clauses. For the indicative mode, the suffix *-ik* in (8) is also used. Perfective constructions, however, do not show the suffix *-aj*, but instead show the suffix *-ej*, which in canonical transitive clauses corresponds to the irrealis mode. These properties are illustrated in the examples (10) and (11).

- (10) [FOC *Leti*] *kíin-s-ik-ech=o'.*
 3.SG die.AF-CAUS-IND-ABS.2SG=CL
 “THAT’S what kills you.” (MDG-B:50)
- (11) [FOC *Leti*] *kíin-s-ej-ø.*
 3.SG die.AF-CAUS-PRF-ABS.3SG
 “HE killed him.” (MDG-B:26)

The best way to describe the agent focus form of the verb in Yucatec Maya is that it constitutes a transitive predicate (which is shown by the specific allomorphs of the mood suffixes of these constructions) that agrees with the object as any other transitive predicate of the language, but in which the bound subject morpheme is absent (the ergative proclitic which is characteristically in cross-reference with the transitive subject). The agent focus form is typically observed when the subject of a transitive clause is focused, but also in transitive subject interrogative and relative clauses.

Finally, focus in Yucatec Maya is subject to a restriction that disallows fronted foci to be nominal expressions introduced by a definite determiner or a demonstrative. This restriction is observed in many other Mayan languages (Aissen, 1992) and is argued in Verhoeven and Skopeteas (2015) to be the result of the interaction between prosodic and syntactic considerations. Example (12) illustrates this restriction. Observe that since the verb shows the agent focus form discussed above, we can be certain that the preverbal subject/agent is a focus and

⁴ See Section 3.1 and Section 4 for a discussion of the different kinds of contrastive focus (corrective, selective, exhaustive, etc.).

not a potentially unmarked subject or some other kind of fronted constituent.

- (12) **Le aj koonol=o' jann-t-ø*
 DET M seller=CL eat.AF-TR-ABS.3SG
oon.
 avocado
 “The seller ate the avocado.” Verhoeven and Skopeteas
 (2015, p. 16)

In short, fronted foci can be unequivocally identified in Yucatec given the robust set of distributional and morphosyntactic properties characteristic of focus in this language.

3 Object fronting and aspect marking in Yucatec Maya

In this section, we present the main word order regularities found in two data sets: a set of 451 VO/OV(S) sentences stemming from a production study conducted in 2017 with 10 Yucatec speakers from Quintana Roo (Section 3.1), and a set of 42 declarative main clauses with fronted direct objects stemming from the 30,000 word long compilation of oral narratives in Yucatec *Maayáaj tsikbalilo'ob Kaampech* (Narraciones Mayas de Campeche, Section 3.2) (Can Canul and Gutiérrez-Bravo, 2016).

3.1 Production study

The elicited data stems from a production study carried out in 2017 with 10 monolingual or Maya-dominant speakers (all female, age range 42–70, mean age 51.9, median age 54) of Yucatec in Yaxley and Tihosuco, Quintana Roo, Mexico. A fieldwork collaborator (bilingual speaker of Yucatec and Yucatecan Spanish) conducted the study in Yucatec and also gave oral instructions to participants that do not read Yucatec. All participants were paid for their participation and gave their informed consent. An AKG C 544L head-mounted microphone connected to a Presonus Audiobox USB was used for recording. Participants were shown a slide presentation in pseudo-randomized order (see [Supplementary material](#)) intended to elicit broad (non-corrective) or contrastive-corrective responses with focus on the direct object. First, participants were shown an example and they were given several suggestions about possible ways to answer. They were then instructed to give only full sentence answers, and to answer in a way that felt natural to them; they were also told that there were no incorrect answers. They were shown two trial examples. Finally, they proceeded to answer the 48 questions of the task proper. The questions for the participants were matched with pictures and accompanying lexical material in order to obtain utterances with three constituents: a subject constituent, an object constituent, and a constituent containing the verb with its tense-mood-aspect (TMA) markers. [Figure 1](#) shows an example of the elicitation of an answer without a correction. The design was $4 \times 3 \times 2 \times 2$, combining four lexicalizations of object constituents with different vowel types (for reasons of prosodic analyses) with three different noun phrases (bare, one modifier, two modifiers),

two focus conditions (non-corrective/corrective), and two vowels (a, i). The non-corrective focus conditions were originally intended to be broad focus cases, but the structure and materials of the elicitation procedure led to contrastive-selective foci.⁵ In the non-corrective conditions, participants were asked “What happens?” (*Ba'ax ku yúuchul?*). A possible answer would be “The woman (*le ko'olelo*) is buying (*ku manik*) a lot of honey (*ya'ab kaab*)”. In the contrastive-corrective focus condition, participants would respond to a biased question such as “The woman is buying many avocados, right?”. A way to respond to this would have been “No, the woman is buying a lot of *honey*”. The subject constituent was always *le ko'olelo* “the woman”, whereas the VP was either *ku jokik* “HAB-ERG.3 uproot” (with the object *xíiw* “grass”) or *ku manik* “HAB-ERG.3 buy” (with all other objects). The objects were: *báat* “ax(es)”, *bak* “meat”, *iib* “beans”, *kaab* “honey”, *kib* “candle(s)”, and *xíiw* “grass”. These were either used in isolation or combined with adjectives (*nukuch* “big”, *chak* “red”, *sak* “white”) or quantifiers (*jach* “very”, *ya'ab* “much”).

The main result is that the 10 Yucatec speakers did not produce a single progressive sentence with a fronted object in either of the two conditions, even though we found abundant instances of each of the two features individually, that is, object fronting ($N = 256$) and progressive marking ($N = 161$), to the degree that the object is predominantly and regularly fronted in habitual sentences. [Table 1](#), [Figure 2](#) show the number of progressive and habitual sentences by word order (VO/OV) for each focus category (Broad/Contrastive): we see that object fronting does not occur in combination with the progressive aspect auxiliary, irrespective of whether the fronting happens to express corrective focus or non-corrective (probably contrastive-selective) focus. In contrast, object fronting occurs predominantly with the HAB marker in the elicited data (see again [Table 1](#), [Figure 2](#)), suggesting that the speakers apparently replace PROG with HAB whenever they intend to focalize the object constituent. Thus, corroborating the observation by [Bohnemeyer](#) (2002, p. 265), our data neatly suggest that there is a strong incompatibility between fronted objects and the progressive (PROG) in Yucatec Maya, which might even be categorical in this language.⁶

3.2 Corpus study

Given the low variability in lexical material included the elicitation study and the very low variability of aspect auxiliaries

⁵ By contrastive-selective, we mean non-exhaustive selection from an open set, or [+identificational] [-contrastive] focus in the sense of [Kiss](#) (1998). Given that the object constituents changed while the verbs remained constant, the broad focus condition still included a contrast between changing nouns in the context of identical actions. The broad focus results might therefore not be fully representative for “out-of-the-blue” sentences, if these can occur in this kind of experiments at all to begin with.

⁶ See footnote 3, and again Section 4, on the question of the (non-)categorical status of this phenomenon. The alternation between OV and VO in the HAB category might be due to the elicitation method, or to some eventual word order flexibility in Yucatec ([Verhoeven and Skopeteas](#), 2015) and will not be fully explored here.

Ba'ax ku yúuchul?

('What happens?')

le ko'olelo'
('the woman')ku manik
('buys')ya'ab kaab
('a lot of honey')

FIGURE 1

Example slide (translations added).

TABLE 1 Progressive and habitual sentences by word order and focus type.

	Broad		Contrastive		Total
	HAB	PROG	HAB	PROG	
VO	22	104	12	57	195
OV	94	0	162	0	256
Total	116	104	174	57	451

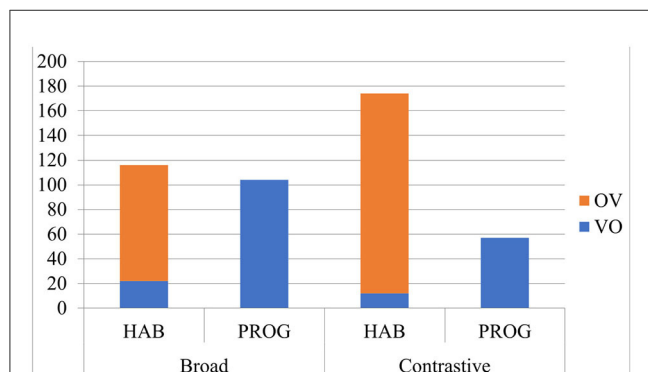


FIGURE 2

Progressive and habitual sentences by word order and focus type.

(only HAB and PROG), we decided to carry out a *post-hoc* corpus study in order to find out if the presumed restriction (no PROG with fronted objects) extends to other types of speech data (i.e., spontaneous speech in oral narratives) and if, apart from the habitual marker, there are other TAM markers (e.g., completive, terminative, present perfect, etc.) that can co-occur with fronted objects.

For this purpose, we extracted all clauses with focus fronting of the direct object that appear in the book *Maayáaj tsikbalilo'ob Kaampech* (Can Canul and Gutiérrez-Bravo, 2016), a compilation of oral narratives in Mayan, 30,012 words in total. We found a total of 42 clauses with fronted focused direct objects: no tokens were found with progressive aspect. The different aspect auxiliaries

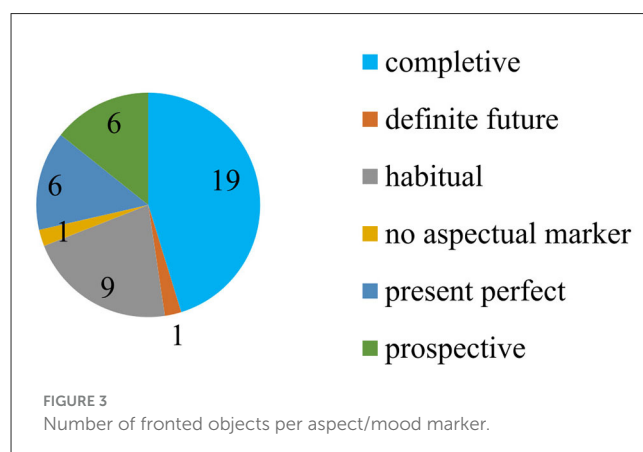


FIGURE 3

Number of fronted objects per aspect/mood marker.

that these focus constructions appeared with are represented in Figure 3.

An additional effect that further supports our observations is the fact that the HAB auxiliary *k* is incompatible with negation, and they never appear together, a well-known fact about Yucatec: instead the progressive is used in this context. However, the prohibition on co-occurrence of the progressive with focus fronting is so strong that it even overcomes this incompatibility. Consider the example in (13).

- (13) *Ma' t=u yutstal u*
 NEG PROG=ERG.2 be.possible ERG.3
y-ok-ol máak=e'.
 EP-enter-IND person=CL
 "It is not possible to go in there." (MTK: 18)
 ("Lit. It is not being possible for a person to enter.")

The text example above refers to a cave that is inaccessible because it is too deep and dangerous. This characteristic of the cave, of course, is a permanent characteristic and consequently one would expect the habitual auxiliary to be used here. However, since the sentential negation and the habitual auxiliary cannot be used together, the progressive auxiliary is used instead. One crucial exception is when a fronted focus or a focus particle (i.e., *chéen*,

“only”) is negated: in this case the co-occurrence of the negation and the habitual auxiliary *is* possible (14).

- (14) *Pero ma'* [FOC *ti'* *tuláakal*] *k=u*
 but NEG PREP all HAB=ERG.3
ts'áab-al=i'.
 give.PASS-IND=CL
 “But they are not given to ALL (people).” (MDG-B: 62)

At first glance, we might interpret the scope of the pre-focal negation as limited to the fronted constituent (Gutiérrez-Bravo, 2015b; Gutiérrez-Bravo et al., 2019). The presence of the clitic *=i'*, however, indicates that, semantically at least, we are still dealing with a case of sentential negation, rendering this observation even more pertinent because it constitutes an exception to the rule that the auxiliary *k* is incompatible with negation. One way to explain this exception is by means of two conflicting constraints, with the restriction on focus fronting together with progressive marking winning over the restriction on negation together with habitual marking. Summing up, what appears to be happening is something along the following lines:

A. The presence of a negation particle with scope over the entire proposition does not allow the presence of the habitual auxiliary, i.e., (13).

B. In case of negated fronted foci, using the progressive would result in a fronted focus construction with progressive aspect, which is also disallowed (Figures 2, 3).

C. Restriction B is stronger and overrides the effects of Restriction A, so in this one single case, negation is allowed in a clause with habitual aspect, i.e., (14).

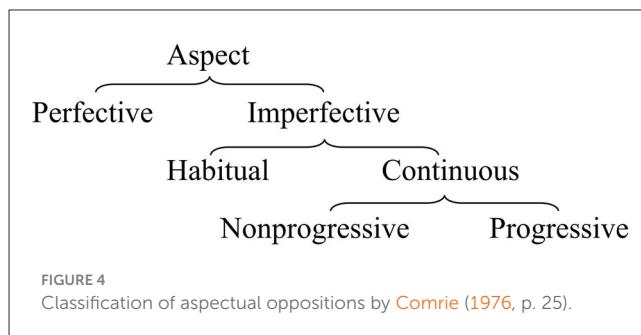
In other words, the prohibition on co-occurrence of the progressive with focus fronting is so strong that it even overcomes a different prohibition, the prohibition on negation plus habitual. This supports our proposal that the progressive is truly disfavored with focus fronting.

Summarizing our main empirical results, we did not obtain a single token of object focus with OV word order co-occurring with the progressive aspectual marker, neither in the production study nor in the compilation of narratives *Maayáaj tsikbalilo'ob Kaampech*. Lastly, we found that there exists a highly unusual construction where the habitual *k* auxiliary co-occurs with the negation particle *ma'*, although the co-occurrence of these two elements is normally banned in Yucatec Maya.

4 Analysis: focus on a singular unbounded event with *táan*

As discussed above, our data suggest that there is a clear-cut incompatibility between object fronting for purposes of focalization and the progressive aspect auxiliary in Yucatec. Our account of this restriction is based on (a) the mereological semantics of habitual and progressive according to Ferreira (2016), and (b) the idea of scalar implicature as a focus related effect according to Rooth (1992).

The habitual auxiliary *k* and the progressive auxiliary *táan* are similar in that both present the event described by the co-occurring verb from an imperfective perspective, meaning they “select an internal part of the target event for inclusion in [the time-reference of the assertion], but no initial or terminal part”



(Bohnenmeyer, 2002, p. 41–42; Smith, 1997). This has been noted in a number of works on the Yucatec TAM-system (Bohnenmeyer, 2002; Vinogradov, 2013; Lehmann, 2017), and it is not an aspectual property specific to Yucatec. Rather, it follows the “most typical subdivisions of imperfectivity” according to Comrie (1976, p. 25) in which the habitual and the progressive are both part of the category of ‘imperfective aspect’ (see Figure 4). Note that such a classification indicates that the imperfective is a more general, overarching category than the progressive and the habitual.

This classification has frequently been adopted and developed further in the formal semantic literature on aspectual categories and event semantics (e.g., Boneh and Jędrzejowski, 2019, pp. 3–5 and references cited therein). According to Ferreira (2016), both progressive and habitual readings are derived from a single imperfectivity operator (*Imp*) that introduces temporal inclusion. Formally (15) states that for any event (description) *e*, the running time of the event $\tau(e)$ is semantically included in the reference (time) interval *i* (Ferreira, 2016, p. 361–362).

$$(15) \text{ Imp} = \lambda P. \lambda i. \exists e: i \subseteq \tau(e) \ \& \ P(e)$$

Such a general semantics for imperfectivity is useful for the analysis of the *k* auxiliary because it is generally considered to be highly grammaticalized and semantically nearly vacuous. Lehmann (2017, p. 203) for example states that today the “auxiliary only survives in its one-phoneme form *k*, obligatorily univertates with the Set A index and carries aspectual information only in contrast with more specific auxiliaries”. In other words, it seems to have undergone semantic bleaching (Sweetser, 1988) and can alternate between a more general, imperfective reading and a more specific, habitual reading, depending on the context. Generalizing sense-shifts from progressives or habituals to imperfect-markers are in fact a frequent cross-linguistic case of semantic bleaching (Bybee et al., 1985; Sweetser, 1988, p. 390).

Now, according to Ferreira (2016), a progressive or a habitual reading arises from the combination between the *Imp* operator and either a singular operator (*sg*) (16) or a plural (*pl*) operator (17). According to (16), there is no event that is a proper subpart of *e*, whereas according to (17), the “*pl* operator takes a set *P* and extracts the homogeneous sums in *P*. These are the sums that can be partitioned into non-overlapping proper parts that are also in *P*” (Ferreira, 2016, p. 358).

$$(16) \text{ sg} = \lambda P. \lambda e. \min(e, P) \\ \min(e, P) \iff P(e) \ \& \ \neg \exists e' < e: P(e')$$

$$(17) \quad \begin{aligned} pl &= \lambda P. \lambda e. \text{sum}(e, P) \\ \text{sum}(e, P) &\iff P(e) \ \& \ \exists e_1, e_2, \dots, e_n < e : \\ &P(e_1) \ \& \ P(e_2) \ \& \ \dots \ \& \ P(e_n) \ \& \\ &\otimes (e_1, e_2, \dots, e_n) \ \& \ e = e_1 \oplus e_2 \oplus \dots \oplus e_n \end{aligned}$$

(Ferreira, 2016, p. 357–358)

When the imperfective operator *Imp* combines with the singular operator ($\text{Imp} \circ \text{SG}$, via function composition), we obtain the progressive reading (18). When it combines with the plural operator ($\text{Imp} \circ \text{PL}$), we obtain the habitual reading (19) (Ferreira, 2016, p. 362).

$$(18) \quad \begin{aligned} [\text{Imp}_{\text{sg}}] &= \text{Imp} \circ \text{SG} = \lambda P. \text{Imp}(\text{SG}(P)) \\ \text{Imp}_{\text{sg}} &= \lambda P. \lambda i. \exists e : i \subseteq \tau(e) \ \& \ \min(e, P) \end{aligned}$$

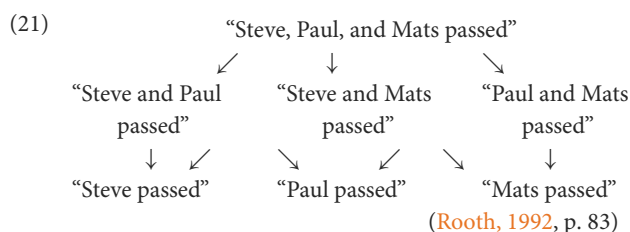
$$(19) \quad \begin{aligned} [\text{Imp}_{\text{pl}}] &= \text{Imp} \circ \text{PL} = \lambda P. \text{Imp}(\text{PL}(P)) \\ \text{Imp}_{\text{pl}} &= \lambda P. \lambda i. \exists e : i \subseteq \tau(e) \ \& \ \text{sum}(e, P) \end{aligned}$$

(Ferreira, 2016, p. 362)

This means that progressive event descriptions refer to *singular* events ongoing at and beyond the reference (time) interval *i*, whereas habitual event descriptions refer to ongoing *sequences* of events that are expected to continue, since the scope of the operator is shifted from the level of one individual event to the level of a number of (recurring) events (Ferreira, 2016, p. 356–357).

Ferreira's account follows Comrie's hierarchy in that progressive and habitual readings are more specific than imperfective readings. As stated above, this is important for the alternation between Yucatec *táan* and *k* because *k* seems to have undergone semantic bleaching and can denote either just temporal inclusion (*Imp*) or combine it with event plurality (*Imp_{pl}*). If we assume that *táan* denotes *Imp_{sg}* and we accept that there is a partial order on sets of events such that asserting the plural occurrence of an event implies a singular occurrence, but not vice versa, we expect *táan* to trigger a scalar implicature in the sense of Rooth (1992). Rooth (1992) shows that, given a set *C* and a partial order \geq_C on *C*, we obtain a pattern in which “asserting an element ϕ of *C* implicates the negation of any higher element of the scale” (Rooth, 1992, p. 82–83). This pattern is illustrated with the example (20) and with the partially ordered set of assertions in (21).

- (20) a. George: How did the test go?
b. Mats: Well, [I]_F passed.



(Rooth, 1992, p. 83)

To describe the effects of this partial order, we need to include groups in the domain of individuals (together with the group sum operation \oplus) and assume that the property *pass* is true of a group *g* exactly if *pass* is true of the atomic parts of *g*. This leaves us with the notation in (22) for *C*. Based on these assumptions, Rooth (1992, p. 83) argues that the assertion “Mats passed” implicates the negation of “Mats and Paul passed”, which in turn implicates that Paul did not pass.

$$(22) \quad \left\{ \begin{array}{l} \text{pass}(s), \text{pass}(m), \text{pass}(p), \\ \text{pass}(s \oplus p), \text{pass}(s \oplus m), \text{pass}(m \oplus p), \\ \text{pass}(s \oplus p \oplus m) \end{array} \right\}$$

(Rooth, 1992, p. 83)

Based on this account, we propose that, following Ferreira (2016, p. 358) and considering a scenario in which three events of “The woman buys a candle” have happened (*e*₁, *e*₂, *e*₃), the denotation of a bare VP such as “The woman -buy- a candle” is (23)a, whereas a progressive reading would be (23)b and a habitual reading would be (23)c. Similar to what has been described by Kiss (1998, p. 266) for object foci, the set of events at which the woman can be said to have bought candles is ordered such that the total set takes the highest value, whereas singular events take the lowest value, with combinations of two events in between.

$$(23) \quad \begin{aligned} \text{a. } [\text{VP}] &= \{e_1, e_2, e_3, e_1 \oplus e_2, e_2 \oplus e_3, e_1 \oplus e_3, e_1 \oplus e_2 \oplus e_3\} \\ \text{b. } \text{sg}([\text{VP}]) &= \{e_1, e_2, e_3\} \\ \text{c. } \text{pl}([\text{VP}]) &= \{e_1 \oplus e_2, e_2 \oplus e_3, e_1 \oplus e_3, e_1 \oplus e_2 \oplus e_3\} \end{aligned}$$

(Ferreira, 2016, p. 358)

Assuming that Rooth's (1992) account of scalar implicatures translates to events, we can predict that marking the singularity of an event via the progressive marker *táan* would implicate the negation of event plurality.⁷ This exhaustivity effect of excluding a set of relevant alternatives is inherently focal. It ranks below corrective focus on the contrastive focus scale in Cruschina (2021, p. 2), presented in (24).

- (24) information focus > exhaustive focus > mirative focus
> corrective focus

This provides us with a straightforward account of the incompatibility between progressive marking and focal object fronting in Yucatec. Specifically, such a combination would contrastively focus both (i) the object to the exclusion of other possible object constituents and (ii) event singularity to the exclusion of event plurality.

In other words, we propose that the progressive in Yucatec is itself a means of focalization, meaning that the auxiliary functions as an operator that selects one particular way of presenting the event. This also means that, at least in Yucatec, the progressive is not entirely restricted to its aspectual function, but, in addition, this aspectual function (presenting the event as ongoing) is also inherently foregrounded against the background of other possible ways of presenting the same event in the relevant discourse situation. A Yucatec sentence with contrastive object fronting and the progressive marker *táan* would therefore trigger two contrastive foci, possibly of different strength. As further discussed below, the corresponding “strategy of inquiry” is probably dispreferred by the speakers since it would target two Questions Under Discussion simultaneously (Beaver et al., 2017, p. 267–269). Therefore, the speakers replace it by the semantically similar, but not focally marked habitual/imperfective *k* auxiliary.

⁷ Note that this implicature would be cancellable (Huang, 2017), but should remain effective until canceled.

5 Discussion

In this section, we discuss the plausibility and validity of our results and the suggested analysis, relating it to both the broader discussion on linguistic evidence and methodologies and other restrictions that have been observed previously in cases of focus fronting.

The most important point to be discussed is our previously mentioned line of argumentation with respect to the question of the (non-)categorical nature of the restriction at hand. It is obviously true that our (categorical) data do not yet show (i.e. prove) that the restriction on the co-occurrence of object fronting and progressive aspect marking is indeed categorical in Yucatec. However, note again that (i) although limited in their scope, our clear-cut results distinctly suggest the potential for a categorical incompatibility between focus fronting and progressive aspect, and (ii) the data predominantly support this hypothesis rather than its opposite, as there is no evidence within our findings that would justify the exploration of an opposing hypothesis. This argument will be referred to as Categorical Hypothesis (CH) in what follows.

Of course, further experimental work is required to show if the restriction that we observe is indeed categorical or rather a dispreferred pragmatic condition/strategy: in the latter case we would of course expect a lower frequency of acceptability when compared to constructions with focus fronting and other aspect auxiliaries. One obvious line of research to verify the CH would be to test the acceptability of fronted focus constructions with PROG in Yucatec with native speakers. However, irrespective of this possibility, our results from two different and independent kinds of data (elicited and corpus data) serve as a valuable point of departure for an investigation of the observed focus effects in Yucatec. Particularly, the fact that progressive marking is relatively frequent in the elicited responses, yet completely absent with fronted objects (under pragmatically very similar conditions), is evidence that we are not dealing with a spurious correlation.

Further note that, even from an acceptability judgement perspective, it is still highly challenging to differentiate whether we are dealing with contextual inappropriateness or straightforward ungrammaticality (Schütze, 2005). Thus, although not suggested by our data, a possible scenario would be one in which object fronting with progressive marking is not strictly ungrammatical, but rather highly inappropriate at the pragmatic level, as one could assume for e.g., (25)c. If we interpret a focused constituent as an answer to a current Question Under Discussion (QUD), two simultaneous foci, on both the fronted object and the singularity of a specific event, would be an attempt to answer two QUDs, such as e.g., (25)a and (25)b, simultaneously. According to Roberts (2012, p. 8), such subquestions should be addressed “one at a time”, meaning that speakers come up with strategies to answer questions by dividing them into subquestions and answering them consecutively one by one (Riester, 2019). This option is not at the participants’ disposal in any kind of straightforward acceptability judgment test of the corresponding phenomenon.

- (25) a. What does the woman buy?
 b. What is the woman doing?
 c. [CANDLES]_F is what the woman is [BUYING]_F.

All in all, we believe that observing different kinds of production data is a necessary and important step in a line of research that needs to encompass future experiments taking into account the caveats we have mentioned. For now, all we can and wish to argue is that our empirical observations form a strong basis for the CH, according to which object fronting is indeed categorically incompatible with progressive marking in Yucatec. Our semantic account renders these observations plausible and provides predictions about the nature of the effect.

Another important point to be addressed is that our observations tie in with those made by Güldemann (2003, p. 323), who elaborates “the hypothesis by Hyman and Watters (1984) that the progressive is an inherently focused verb category”. He observes that marking mechanisms for predication focus and present progressive frequently show a formal similarity in Bantu languages. In some instances, this isomorphism can be seen as a consequence of a directional grammaticalization shift, moving from predication focus to progressive. Güldemann (2003, pp. 349–350) argues that “a present progressive can focus on the HIC-ET-NUNC of the relevant state of affairs [and] can also enhance another semantic component, namely the ongoing, continuous nature of the event. [...] This in turn is closely related to the notion of emphasis and builds another possible bridge between the progressive and the focus of an utterance”. Similarly, Smith (1997, p. 74) states that “the progressive viewpoint has meanings that do not arise for other types of imperfective. Nuances of activity, dynamism, and vividness are often associated with the sentences of this viewpoint.” We consider these observations to be pragmatic extensions of the semantics laid out in Section 4.

It is also worth commenting that restrictions on fronted foci (syntactic, semantic, prosodic, or a combination thereof) are far from being rare. For instance, Pusch (2003) and Cruschina and Remberger (2017) observe that focus fronting in Romance languages is incompatible with negation, and with imperative and other markers of illocutionary force. Cruschina and Remberger (2017) further observe that, in most Romance languages, nominal and adjectival predicates of copular constructions cannot be fronted foci. Also, focus fronting is largely restricted to contrastive foci in Romance languages (Zubizarreta, 1998; Cruschina and Remberger, 2017) and in Hungarian (Kiss, 1998).⁸ Finally, recall from Section 2 that Yucatec itself has previously identified restrictions on fronted foci (e.g., the exclusion of nominal expressions introduced by the definite article). As such, the restriction on fronted foci that we argue for and document here for Yucatec should not be considered a peculiar quirk. Rather, it adds to the long list of restrictions observed cross-linguistically on focus fronting.

6 Conclusions

In this article, we have presented evidence of a restriction on focus structures in Yucatec Maya that makes object fronting incompatible with the progressive aspect marker *táan*. In our production study, object fronting and progressive aspect marking are mutually exclusive, even though we found abundant instances

of each of the two individually ($N_{OV} = 256/N_{PROG} = 161$). Our *post-hoc* corpus study points in the same direction, since none of the 42 clauses with fronted focused direct objects in *Maayáaj tsikbalilo'ob Kaampech* (30,012 words in total) showed progressive aspect. In order to account for the observed incompatibility in our data, we have proposed that the (use of the) progressive marker implies a contrastively focal function, which we understand as asserting event singularity (following Ferreira, 2016) to the exclusion of event plurality. This exclusion arises as a scalar implicature in the sense of Rooth (1992) and would lead to a double focus construction if combined with a fronted object.

We also discussed whether the observed incompatibility should be understood as a case of ungrammaticality or rather as a sub-optimal way of information packaging relative to the QUD structure of a given discourse. Although solving this particular issue goes beyond the scope of this work, we have nonetheless presented suggestions for future research regarding possible intervening factors, while at the same time arguing for the validity of production data as a point of departure for such an investigation. Finally, we argued that our observations should be connected to the isomorphism between marking mechanisms for predication focus and present progressive in Bantu languages, as well as to restrictions on fronted foci in Romance languages, again lending plausibility to the idea that the progressive auxiliary in Yucatec Maya is focal and therefore disallowed in sentences with fronted focused objects.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

8 Cruschina (2006) reports fronted informational foci in Sicilian, meaning that Sicilian apparently does not limit focus fronting to contrastive contexts.

References

- Aissen, J. (1992). Topic and focus in Mayan. *Language* 68, 43–80. doi: 10.1353/lan.1992.0017
- Beaver, D. I., Roberts, C., Simons, M., and Tonhauser, J. (2017). Questions under discussion: where information structure meets projective content. *Ann. Rev. Linguist.* 3, 265–284. doi: 10.1146/annurev-linguistics-011516-033952
- Bohnemeyer, J. (2002). *The grammar of time reference in Yucatec Maya*. LINCOM Studies in Native American linguistics: Vol. 44. München: LINCOM.
- Boneh, N., and Jędrzejowski, Ł. (2019). Reflections on habituality across other grammatical categories. *STUF – Lang. Typol. Universals* 72, 1–20. doi: 10.1515/stuf-2019-0001
- Bricker, V. (1979). “Wh-questions, relativization and clefting in Yucatec Maya,” in *Papers in Mayan Linguistics*, ed L. Martin (Columbia, Missouri: Lucas Brothers), 107–136.
- Bricker, V., Po'ot Yah, E., and Dzúl de Po'ot, O. (1998). *A Dictionary of the Maya Language as Spoken in Hocabá, Yucatán*. Salt Lake City: University of Utah Press.
- Bybee, J. L., Pagliuca, W., and Dressler, W. U. (1985). “Cross-linguistic comparison and the development of grammatical meaning Suppletion in word-formation,” in *Historical Semantics - Historical Word-Formation*, ed. J. Fisiak (Boston: De Gruyter Mouton), 59–84.

Author contributions

MU: Formal analysis, Visualization, Writing—original draft, Writing—review and editing, Conceptualization, Data curation, Investigation, Methodology, Project administration, Resources, Supervision, Funding acquisition, Validation. RG-B: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Writing—original draft, Writing—review and editing. JF: Formal analysis, Project administration, Visualization, Writing—original draft, Writing—review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Project numbers 281511265 (DFG project A5, CRC 1252 at University of Cologne) and 491466077 (DFG open access publication fund at University of Potsdam).

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/flang.2023.1286520/full#supplementary-material>

- Can Canul, C., and Gutiérrez-Bravo, R. (2016). *Maayáaj tsikbalilo'ob Kaampech: Narraciones mayas de Campeche*. Mexico: Instituto Nacional de Lenguas Indígenas.
- CDAM (1984). *Alfabeto maya acordado en la reunión de agosto de 1984*. Mérida: INEA/Academia de la Lengua Maya de Yucatán.
- Comrie, B. (1976). *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems* (Cambridge textbooks linguistics). Cambridge, New York: Cambridge University Press.
- Cruschina, S. (2006). "Informational focus in Sicilian and the left periphery," in *Phases of Interpretation*, ed. M. Frascarelli (Berlin: Mouton de Gruyter), 363–386.
- Cruschina, S. (2021). The greater the contrast, the greater the potential: on the effects of focus in syntax. *Glossa* 6, 1–30.
- Cruschina, S., and Remberger, E. (2017). "Focus fronting," in *Manual of Romance Morphosyntax and Syntax*, eds. A. Dufter and E. Stark (Berlin, Boston: De Gruyter), 502–535.
- Ferreira, M. (2016). The semantic ingredients of imperfectivity in progressives, habituals, and counterfactuals. *Nat. Lang. Semant.* 24, 353–397. doi: 10.1007/s11050-016-9127-2
- Güldemann, T. (2003). Present progressive vis-à-vis predication focus in Bantu. *Stud. Lang.* 27, 323–360. doi: 10.1075/sl.27.2.05gul
- Gutiérrez-Bravo, R. (2015a). "El foco de verbo en maya yucateco," in *Nuevos estudios de la Teoría de la Optinidad*, eds. R. Gutiérrez-Bravo, F. Arellanes Arellanes and M. Chávez Peón (Mexico: El Colegio de México), 17–48.
- Gutiérrez-Bravo, R. (2015b). *Las cláusulas relativas en maya yucateco*. México, D.F: El Colegio de Mexico.
- Gutiérrez-Bravo, R. (2017). Clefts and focus in Yucatec Maya. *Cuadernos de Lingüística de El Colegio de México* 4, 5–47. doi: 10.24201/clecm.v4i1.52
- Gutiérrez-Bravo, R., and Monforte, J. (2010). "On the nature of word order in Yucatec Maya," in *Information Structure in Languages of the Americas*, eds. J. Camacho, R. Gutiérrez-Bravo, and L. Sánchez (Berlin/New York: Mouton de Gruyter), 139–170.
- Gutiérrez-Bravo, R., and Monforte, J. (2011). "Focus, agent focus and relative clauses in Yucatec Maya," in *New Perspectives in Mayan Linguistics*, ed. H. Avelino (Cambridge: Cambridge Scholars Publishing), 257–274.
- Gutiérrez-Bravo, R., Sobrino, M., and Uth, M. (2019). "Contrastive focus in Yucatecan Spanish," in *The Syntactic Variation of Spanish Dialects*, ed. A. J. Gallego. Oxford: Oxford University Press, 275–296.
- Huang, Y. (2017). "Conversational implicature," in *Oxford Handbooks in Linguistics. The Oxford Handbook of Pragmatics*, ed. Y. Huang (Oxford: Oxford University Press), 156–179.
- Hyman, L. M., and Watters, J. R. (1984). Auxiliary focus. *Stud. Afri. Linguist.* 15, 233–273. doi: 10.32473/sal.v15i3.107511
- Kiss, K. É. (1998). Identificational focus versus information focus. *Language* 74, 245–273. doi: 10.1353/lan.1998.0211
- Lehmann, C. (2008). "Information structure and grammaticalization," in *Typological Studies in Language: v. 77. Theoretical and Empirical Issues in Grammaticalization*, ed. E. Seoane and M. J. López-Couso (Amsterdam/Philadelphia: John Benjamins Publishing Company), 207–230.
- Lehmann, C. (2017). "Grammaticalization of tense/aspect/mood marking in Yucatec Maya," in *Unity and Diversity in Grammaticalization Scenarios*, eds. A. Malchukov and W. Bisang (Erscheinungsort nicht ermittelbar: Language Science Press), 173–237.
- Pusch, C. D. (2003). "Aspectuality and Focality-Reflections on semantics-pragmatics relations and isomorphism in Romance progressive periphrases," in *Romanistik in Geschichte und Gegenwart: Vol. 9. Verbalperiphrasen in den (ibero-)romanischen Sprachen*, ed. C. D. Pusch (Hamburg: Buske), 179–192.
- Riester, A. (2019). "Constructing QUD trees," in *Current Research in the Semantics/Pragmatics Interface: Volume 36. Questions in Discourse. Volume 2: Pragmatics*, eds M. Zimmermann, K. v. Heusinger, and E. Onea Gáspár. (Leiden, Boston: Brill), 164–192.
- Roberts, C. (2012). Information structure in discourse: towards an integrated formal theory of pragmatics. *Semant. Pragmat.* 5, 1–69. doi: 10.3765/sp.5.6
- Rooth, M. (1992). A theory of focus interpretation. *Nat. Lang. Semantics* 1, 75–116. doi: 10.1007/BF02342617
- Schütze, C. T. (2005). "Thinking about what we are asking speakers to do," in *Studies in Generative Grammar: Vol. 85. Linguistic Evidence. Empirical, Theoretical, and Computational Perspectives*, eds. S. Kepser and M. Reis (Berlin: Mouton de Gruyter), 457–484.
- Skopeteas, S., and Verhoeven, E. (2012). "Left-peripheral arguments and discourse interface strategies in Yucatec Maya," in *Contrasts and Positions in Information Structure*, eds. I. Kučerová and A. Neeleman (Cambridge and Mass: Cambridge University Press), 296–321.
- Smith, C. S. (1997). *The parameter of aspect. Studies in Linguistics and Philosophy: Vol. 43*. Dordrecht; London: Kluwer Academic.
- Sweetser, E. E. (1988). Grammaticalization and semantic bleaching. *Ann. Meet. Berkeley Linguist. Soc.* 14, 389–405. doi: 10.3765/bls.v14i0.1774
- Tonhauer, J. (2003). "F-constructions in Yucatec Maya," in *The Proceedings of SULA 2*, eds. J. Anderssen, P. Menéndez-Benito, and A. Werle (Amherst: GLSA, University of Massachusetts), 203–223.
- Verhoeven, E., and Skopeteas, S. (2015). Licensing Focus Constructions In Yucatec Maya. *Int. J. Am. Linguis.* 81, 1–40. doi: 10.1086/679041
- Vinogradov, I. (2013). El aspecto progresivo en las lenguas mayas. *Anales de Antropología* 47, 127–154. doi: 10.1016/S0185-1225(13)71021-X
- Zubizarreta, M. L. (1998). *Prosody, focus, and word order. Linguistic Inquiry Monographs: Vol. 33* (Cambridge, Mass.: MIT Press).



OPEN ACCESS

EDITED BY

Joseph V. Casillas,
Rutgers, The State University of New Jersey,
United States

REVIEWED BY

Saidna Zulfiqar Bin Tahir,
Universitas Iqra Buru, Indonesia
Hannah Treadway,
University of Florida, United States
Muassomah,
Universitas Islam Negeri Maulana Malik
Ibrahim, Indonesia

*CORRESPONDENCE

Abdullah M. Seraye Alseraye
✉ aalseraye@ksu.edu.sa

RECEIVED 04 November 2023

ACCEPTED 16 February 2024

PUBLISHED 08 March 2024

CITATION

Alseraye AMS (2024) L2 Arabic learners' processing of Arabic garden-path sentences: a consistent reading pattern.
Front. Psychol. 15:1333112.
doi: 10.3389/fpsyg.2024.1333112

COPYRIGHT

© 2024 Alseraye. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

L2 Arabic learners' processing of Arabic garden-path sentences: a consistent reading pattern

Abdullah M. Seraye Alseraye *

Department of Curriculum & Instruction, School of Education, King Saud University, Riyadh, Saudi Arabia

Purpose: The main purpose of this study was to investigate to what extent the L2 Arabic learners' reading process is affected by the incomplete representation of speech (the absence of short vowels and diacritics) while reading ambiguous sentences (garden path sentences).

Method: With a self-paced reading software program, 41 non-native male students, aged from 22 to 26, enrolled in King Saud University, participated in reading 44 sentences (followed by reading comprehension questions) representing three reading conditions, plain, vowelized-discretized, and wrongly-vowelized.

Results: For the reading times data, the analysis revealed that the GP structure had a significant effect on the reading processes of L2 Arabic learners; it took them longer to read the GP sentences than their non-GP counterparts. For the reading comprehension, the analysis did not reveal any significant differences between the means for the percentages of correct responses. For the comparison between the three reading conditions, a significant difference was found: it took the participants on average less time to read the GP sentences when presented plain, and more time with the incorrect representation. However, their reading comprehension was not affected.

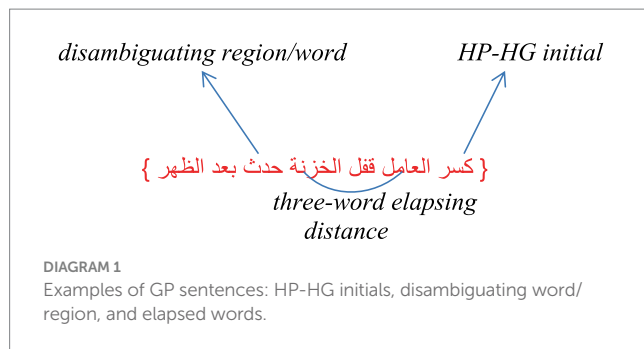
Conclusion: In addition to the good-enough model and the nature of Arabic morphology, the reading experience, is a good candidate to start with as an important factor in the interpretation of the ineffectiveness of the GP structure on the reading comprehension process of Arabic readers, in which the segregability of Arabic writing system prepare the readers to emphasize some sensory inputs and ignore others based on their past reading experience.

KEYWORDS

garden-path structure, Arabic short vowels, past experience, heterophonic-homographic initial, L2 Arabic learners, reading comprehension

Introduction

Languages are intrinsically susceptible to structural ambiguity. Indeed, "at any given point in a sentence, the available information can be ambiguous at many levels," because languages are "structured at multiple levels simultaneously, including lexical, phonological, morphological, syntactic, and text or discourse levels" (MacDonald et al., 1994, p. 676). This structural ambiguity can result from either an optional, controlled cause or a non-optional, compulsory cause. Some researchers view the Arabic writing system as being by nature



ambiguous due to the consonantal representation of its orthography (De Francis, 1989). However, structural ambiguity in Arabic is not omnipresent and internally structured but situational because of its optional and segregable representation of short vowels and diacritics.

In fact, integrating the necessary short vowels and diacritics with the consonants would turn Arabic print into a transparent orthography and should therefore disambiguate any potential structural ambiguity in the sentence by leaving only one acceptable reading for each word. Conversely, the absence of the necessary short vowels and diacritics from the consonantal representation in Arabic print would turn its orthography into a deep orthography, which means highly ambiguous orthography. An extreme example of deep orthography is a sentence that begins with a heterophonic homographic word (henceforth: HP-HG); this becomes much worse when the disambiguating region of the sentence is far from its initial word, which would very likely give rise to the garden-path phenomenon (GP) (Seraye, 2004; Alseraye, 2022). This farness in distance from the disambiguating region has been found to be from three to five words distance (Seraye, 2004, 2016; Alseraye, 2022).

In the modern Arabic language, the order of the words in a sentence is flexible, and it may take, based on stylistic variations, either one of the following word orders: VSO (i.e., verb, subject, object), SVO, VOS, and OVS (Mohammad, 2000), with no syntactic preference for one order type over the others. On the surface, this first word of the sentence could be a verb phrase (VP), as in “كسر العامل قفل الخزانة,” a noun phrase (NP) as in “العامل كسر قفل الخزانة,” or a prepositional phrase (PP), as in “في الخزانة مال.” In these word orders, the ambiguity is expected with a sentence order that begins with a VP that has a third person singular verb in the past tense with enough words go in between the initial HP-HG word and its disambiguating region. For an illustration, see Diagram 1.

In the following, a descriptive syntactic analysis of a GP Arabic sentence is laid out:

كسر العامل قفل الخزانة حدث بعد الظهر [كسر العامل] [1] ... كسر العامل

Reading 1:

- *The worker's breaking of the lock of the safe [occurred] in the afternoon* (grammatically correct).
- (a) Arabic word order: breaking + the worker + the lock + of the safe + occurred + in + the afternoon.
Correct structure: S [NP [NP, NP, PP[P, NP]], VP[PP[NP]]]
- (b) English word order: the worker's breaking of the lock of the safe occurred in the afternoon.

Reading 2:

- *{broke} the worker the lock of the safe [occurred] in the afternoon* (grammatically incorrect).

(a) Arabic word order: broke + the worker + the lock + of the safe + occurred + in + the afternoon.

Mistaken structure: S [VP [NP, NP, PP [P, NP]], VP [PP [P, NP]]].

(b) English word order: the worker broke the lock of the safe occurred in the afternoon (Alseraye, 2022).

This HP-HG VP has numerous forms of the same verb that still share the same semantic root but also reflect the following different case roles in the sentence: an active verb, a gerund, and, to some degree, a passive verb in cases where the writer/speaker wants the importance or focus to be shifted toward who/what experiences the action. However, there are three conditions in which the structural ambiguity would, at its high peak, lead readers astray by garden pathing them. The first condition is when the initial word of the sentence is an HP-HG word that has numerous forms, while for the second the initial word is presented as plain, such as without the right short vowels and diacritics (i.e., unvowelized and undiacritized). For the third condition, an adequate distance elapses between the initial HP-HG word and its disambiguating region to prevent the eye from visualizing the subject and the predicate simultaneously. Therefore, these two essential parts of the sentence are not in the reader's peripheral vision simultaneously (Diagram 1).

In the second condition, the distance that elapses between the initial HP-HG word of the sentence and its region of disambiguation is enough [five words with skilled adult readers, such as in Alseraye (2004) study; three words with the children who are beginning readers, as in Alseraye (2022) study]. This is because it helps to prevent the parser from benefiting from the linguistic context; that is, from the neighboring words that would help the parser, to a large extent, to recognize the correct form, build the right “structure tree,” and hence, avoid the disturbance expected at the region of disambiguation. Using eye movement terminology, this occurs when the reader's eye does not capture, at one fixation, both essential parts of the sentence, the subject, and its predicate, within one visual span.

Even using a clear example of an ambiguous sentence (a passive sentence where the initial word is a verb in the passive voice), figuring out whether the verb is active or passive can sometimes be achieved without needing to reach the region of disambiguation to get it right, and without the reader being garden pathed. This can be figured out using different sources, such as the immediate previous context, the text discourse, and the reader's experience exemplified in word and syntactic structure frequencies. Indeed, we would not expect to find such a great distance elapsing between an initial verb in the passive voice and its subject that this would garden path the reader, as are exemplified in the sentences used in Roman et al. (1985) and Hermena et al. (2015) studies. The exceptions occur in artificially created sentences. According to Ots (2021), the speaker, in the linguistic encoding stage in language production, would “assign the syntactic functions [that] are appropriate for the message and order the constituents, given the discourse and grammatical constraints” (p. 2). Similarly, the writer, I assume, would be forced to comply with these constraints, in addition to the cognitive constraints and limitations, to avoid the long distance between the HP-HG word and its disambiguating region in the sentence. The two sentences below

illustrate the GP sentence as used by Hermena et al. (2015) (labeled “1”) and the modification of the same sentence that a writer is expected to make to avoid the ambiguity of the passive structure (labeled “2”). There is also a basic naturally constructed GP sentence (labeled “3”):

1. سمع صرخة مدوية حينما دفعت1 الطالبة التي كانت في طريقها إلى المعمل بيدي2 زميلتها3 فهوت م.

Translation: “Everyone heard a loud scream when the student pushed (or was pushed) on her way to the lab (by the hand of) her colleague so she fell unconscious.”

2. سمع الجميع صرخة مدوية حينما تم دفع1 الطالبة التي كانت في طريقها إلى المعمل بيدي2 زميلتها3

3. كتب محمد منسية في مدرسته.

Translation: Mohammed’s books were forgotten/left out in his schools.

The structure of the clause in sentence 1 includes a passive voice, and without the supplement of the correct short vowels and diacritics the readers would be garden pathed. They would, as demonstrated by previous studies (Al-Fahid, 2000; Hermena et al., 2015; Seraye, 2004, 2016; Alseraye, 2022), read the HP-HG verb “دفع” as an active verb “(she) pushed: دَفَعَتْ” Then, when they reached the disambiguating region, such as the word “by the hand of” بيدي they would notice that their decisions were wrong, because they did not read the verb “دفع” as being in the passive voice (seven words elapsing). However, in practical terms, it is expected that when writers express the meaning of such a sentence or clause they would not generally include a clause with such an ambiguous passive voice in a sentence that is so complex. They would avoid this by converting the verb “دفع” into a phrase that says the same thing without any need to refer to an explicit subject. At the same time, they can eliminate the ambiguity without having to vowelize and diacritize the verb, by using the phrase “تم دفعها” Another strategy to avoid sentence ambiguity is to use a syntactic mechanism called topicalization, in which the focus is brought to the front, and the order of items, such as the constituents, the phrases, and its adjuncts, changed so they are close to the main verb.

The example of the GP sentences (number 3) used in Alseraye (2022) is, to some degree, representative of the type of sentences that Arab children might see in a text that is realistic and naturally structured. However, Arab readers were nevertheless garden pathed, unless they were able to guess (see Appendix A).

In addition to the current study, previous studies (Seraye, 2004; Alseraye, 2022) have already examined the behaviors of the Arab readers, both beginners (the children) and experienced (the adults), when they read. The studies have found that Arab readers who read ambiguous sentences, such as the GP ones that were embedded in short texts for reading aloud, demonstrated reading behaviors that were essentially consistent among all participants regardless of their reading skills and reading-equivalent ages (i.e., experience). Those reading behaviors are as follows:

- Very frequently, when the basic verb is embedded in a discourse, such as a text, it does matter whether the verb was correctly vowelized and diacritized. The Arabic learners will read it as an active voice. To illustrate, the participants read the HP-HG verbs

in the passive voice as “أعلن was announced” and “سمعت was heard,” as/أَعْلَنَ [he] announced / and / سَمِعَتْ [she] heard/ (i.e., active voice form), although the linguistic context would force them to read the two verbs as passive voice forms.

- The verbs in the passive voice, “أُشْتُهِرْتُ was known” and “أُرْتُكِبْتُ was committed,” were read as active voice forms, although they were provided with the right short vowels and diacritics that would make them non-homographic words.
- Although the participants paused and hesitated over some of the HP-HG words, they were garden pathed and very frequently made no regression (reanalysis) by going back to their initial decisions and choosing the right forms of the HP-HG verbs. Indeed, some of the participants would not even pause over the HP-HP words.
- Some participants, once they realized that the forms were passive, applied the knowledge they had acquired to the verbs that followed, by initially considering them as passive forms. That is, their previous experience with the first passive verb they encounter acts as a prime and forces them to read the first verb that follows as a passive one. However, when the participants found out that they had been garden pathed, some of them made exclamations such as “لا!” meaning “No!”
- Some of the participants insisted on using the active voice form for the vowelized/diacritized passive voice “أُرْتُكِبْتُ” pausing over the verb, saying it as an active form, although it was represented as a passive voice form.
- When the HP-HG verb, in its plain condition, had no previous context to force the reader to read it as passive rather than active, it was read as an active form. Indeed, even when the basic verb was provided with the right short vowels and diacritics that, if assembled with the consonants, would be read as a gerund, participants read it as being active, with the short vowels and diacritics ignored, and were thus garden pathed.
- An automatic attempt was made to convert a verbal noun or a gerund (e.g., اسْتَنَّاف appealing) into an active basic verb (استأنف [he] appealed).

In conclusion, our study, based on the post criterion assessment (reading aloud task) found the same phenomenon among the learners of Arabic as a second language (L2) who are at the advanced level and qualified to enroll in academic programs where the language of communication is Arabic as was observed within native Arab readers, both children and adults.

Therefore, we can conclude as Seraye (2004) has already stated that “it seems that the initial sentence default, to use the notions of the symbolic and associative theories of cognition (Marcus et al., 1995), was the verb and not the noun or the preposition which Arabic allows. Further, this default was characterized by the fact that it was always regarding an active-voice verb, and this was noticed even in an embedded clause when the sentence led the reader logically to a passive voice more than to an active voice” (pp. 135–136).

In the literature, the ambiguity associated with the GP structure drew the attention of researchers, who set up suitable apparatuses to answer the following two questions: “[How do] people cope with rampant ambiguity, especially syntactic ambiguity, as the linguistic signal unfolds over time? [H]ow is sentence interpretation affected by variations in syntactic complexity?” (MacDonald and Hsiao, 2018, p. 173; Alseraye, 2022).

In the Arabic literature, research has focused, to a large extent, on the text and word levels, and very few studies have addressed the structural ambiguity, particularly the GP sentences, that is caused by the absence of short vowels and diacritics (for an overview, see Roman et al., 1985 [in French, and cited in Hermena et al. (2015)]; (Seraye, 2004, 2016; Ibrahim, 2013; Taha, 2016; Saiegh-Haddad, 2017; Abu Rabia, 2019; Hermena and Reichle, 2020; Alseraye, 2022)).

In the syntactic ambiguity that has been addressed recently is ambiguity that is caused by the GP structure in Arabic and results from the optional, segregable nature of Arabic orthography. The current survey of the studies conducted in Arabic shows that very few that have been published in the literature have addressed the processing of structural ambiguity caused by the GP structure and its effect on reading accuracy and comprehension.

A study by Abu-Rabia (1995) claimed that Arabic readers can not read or comprehend sentences that were not properly vowelized unless they reanalyzed their first readings of the initial HP-HG head of the ambiguous sentence. In response to this assumption, Seraye (2004) conducted a second experiment that assessed, among other variables, the effect of the GP structure on the reading processes, reading times, and comprehension of highly skilled Arab adults ($n = 35$, in the 26–40 age range), in correlation with the presence and absence of correct short vowels and diacritics.

Using a self-paced reading, controlled by a moving window software program in which the sentence is read word-by-word without regression, Seraye found that the Arab adults' reading times of the GP sentences were affected ($p = 0.016$), but that their reading comprehension was not ($p = 0.053$). The average length of time to read the two types of sentences was longer for the GP sentences ($M = 6,747.14$ ms) than for the non-GP ones ($M = 6,259.30$ ms), but their reading comprehension performance was very good on both types of sentences ($M = 0.89$, for the GP sentences, and $M = 0.83$, for the non-GP sentences, $SD = 0.08$). Therefore, from a descriptive perspective, the mean values indicate a positive relationship between the reading time length of the GP/non-GP plain sentences and the reading comprehension performance.

In addition, Alseraye's (2004) third experiment on word naming using the E-Prime software program found that reading latency was "positively correlated with the gradual increase of the number of short vowels and diacritics" in comparison to the consonants (p. 214).

This ineffectiveness of the GP structure regarding the reading comprehension of Arab adults was attributed to two features of Arabic morphology: the core semantic element, based on the *trilateral/quadrilateral*-root, that is shared among all activated forms of HP-HG word, and the form/pattern of the Arabic word, its skeletal tier/word pattern/binyan (McCarthy, 1979, 1981), which would narrow the possible readings of the HP-HG word (Seraye, 2004). In fact, as proposed by Seraye (2004), "the predictability/productivity of word forms/patterns, affixation, etc., compensate for the lack of short vowels and diacritics in print" (p. 259).

This advantage of Arabic morphology roots and word patterns in the Arabic reading process has been, consolidated, and theoretically grounded (Seraye, 2004, 2016; Mahfoudhi, 2007; Mahfoudhi et al., 2010; Abu-Rabia, 2012; Abu-Rabia and Abu-Rahmoun, 2012; Boudelaa and Marslen-Wilson, 2015; Taha and Saiegh-Haddad, 2016, 2017; Maroun, 2017; Saiegh-Haddad, 2017; Abu Rabia, 2019; Aljasser, 2020; El Akiki and Content, 2020; Hermena and Reichle, 2020; Wattad and Abu-Rabia, 2020; Abu-Rabia, 2021; Alseraye, 2022; Khateb et al., 2022; Aldholmi and Pycha, 2023).

The question became whether adding appropriate short vowels and/or diacritics to the initial HP-HG words would help in blocking the GP phenomenon and enhance the reading process by minimizing reading times and speeding up the parser's checking processes. In response to the question, Seraye (2004) compared four reading conditions: plain (rc1), short vowels-plus-shaddah (rc2), sukun-only (rc3), and case-ending marking-only (rc4).

The analyses showed no significant results regarding reading conditions for either reading time ($p = 0.283$) or reading comprehension ($p = 0.237$). Examining the total means visually shows that the participants took more time to read rc2 ($M = 7,277.76$) and rc4 ($M = 7,230.64$), and less time to read rc1 ($M = 6,747.14$). For their performance on reading comprehension questions, the percentages of their correct answers were on average very good despite the reading condition (the correct answers percentages range is between 0.89, for rc1, and 0.80, for rc2).

Using the eye movement technique, and with special types of sentences (passive voice), Hermena et al. (2015) examined the effect of Arabic orthographic representation on the reading processes of 25 adult native Arabic speakers, collecting eye movements measures/data on different regions of the sentence. Five reading conditions were constructed by manipulating either the initial HP-HG word or the entire clause, that is embedded in a very complex structural sentence, as exemplified and illustrated above. Only when the initial HP-HG word of the clause was passive and presented as plain would the reader be garden-pathed.

Among the findings revealed by Hermena et al.'s (2015) study, is that Arab adults takes more time to read GP sentences (embedded clause), and that their reading comprehension was not affected (the correct answers percentages range is between 70 and 100%). In addition, a longer fixation duration was observed on the disambiguating region of the GP sentence once the HP-HG initial of the GP sentence was a passive verb and presented as plain.

Since the population in previous studies included highly skilled adult readers, the explanation that relates to experience in the previous findings would garner more support if less experienced Arab readers were incorporated. Therefore, the target sample in Alseraye's (2022) follow-up study was beginning Arab readers. A total of 39 fourth-grade native Arabic speakers, at the age of 9–10, were included. With the same self-paced moving window software program used by Seraye (2004), the participants read 36 actual seven-word sentences (of both, GP and non-GP sentences) and eight practice sentences representing three reading conditions. These included a plain condition in which only the consonants were presented, a fully vowelized and diacriticized condition, and an incorrectly vowelized condition by manipulating the short vowels only incorrectly while keeping the consonants intact. After reading each sentence, a comprehension question would pop up with three response options: *true*, *false*, and *I do not know*.

The analysis did not reveal any significant differences between the GP and non-GP sentences on reading times ($p = 0.710$) or reading comprehension ($p = 0.105$).

However, examining the overall means showed that it took the participants longer on average to read the GP sentences ($M = 8,172$ ms) than the non-GP ones ($M = 8,113$ ms). For reading comprehension, the overall means for the non-GP and GP sentences were $M = 0.73$ and $M = 0.80$, respectively. Furthermore, when the GP sentences (in the plain condition) were compared to the other two reading conditions, no significant results were found, indicating that the participants'

reading times were the same on average ($p = 0.565$). However, examining the overall mean values showed the following: it took the participants 8,172.33 ms on average to read the GP plain sentences, 8,007.64 ms to read the GP vowelized and diacritized sentences, and 7,882.74 ms to read the GP wrongly vowelized sentences. The overall means showed that it took the participants more time on average to read the plain sentences than the vowelized-diacritized ones. However, the participants benefited from the presence of the short vowels and diacritics that resolved the GP structure; they took less time to read these in comparison to the plain reading condition. Since they took far less time to read the GP sentences that included incorrect short vowels and diacritics, this should have had no effect on their reading time as in the plain reading condition. However, this was not the case. The findings were objectivized as equivocal findings, and the only trend that could be extracted from the results regarding the reading times is that the GP sentences in plain representation took the Arab readers longer to process, regardless of their reading levels (skilled versus beginning). This is consistent with previous findings.

On the other hand, the analysis of the data on reading comprehension revealed a significant difference between the three reading conditions ($p = 0.026$). Pairwise comparisons showed significant differences between reading condition 1 and reading condition 2, ($p = 0.045$) and reading condition 1 and reading condition 3 ($p = 0.012$). However, there was no significant difference between reading condition 2 and reading condition 3 ($p = 0.618$). Examining the means values, however, shows that the participants scored higher on average on condition 1 ($M = 0.80$) than on condition 2 ($M = 0.69$) and condition 3 ($M = 0.66$). This finding that the participants understood the GP plain sentences better than their counterparts in the other reading conditions, is consistent with previous studies (Seraye, 2004; Alseraye, 2022). The trend noted from the previous studies on reading comprehension is that on average the participants understood the GP sentences better than their non-GP counterparts, and that a correlation, from a descriptive perspective, could be inferred visually between the reading times and reading comprehensions; that is, the more time readers of Arabic spend reading the GP, the more accurate responses they score.

In conclusion, the reading behaviors of Arab adults and children regarding the GP sentences showed the following: the persistence of initiating the active basic form of each HP-HG word by making it a default despite its orthographical representation; an automatic attempt to convert the gerund into an active basic verb; and finally, ignoring the supplemental short vowels and diacritics. Furthermore, the statistical results of the previous studies on the GP sentences showed that Arab readers, both adults and children, can read and comprehend the print even if it is presented incompletely, and that they do not need to process the GP sentences twice to comprehend them. Taking the findings of the descriptive and statistical data together leads to questions about the characteristics of the Arabic parser, particularly in terms of the apparatuses that are relied upon in analyzing a consonantal representation of Arabic.

The only factor that can still be suggested as being implicated in the processing of GP sentences in Seraye (2004) and Alseraye (2022) is the reading experience. Therefore, there is likely to be a factor that is involved with and precedes the visual processing of print and that interferes automatically with the visual processing of print even when the writing systems used do not represent speech accurately and completely by vowelizing and diacritizing. It is suggested that this

factor is the reader's previous exposure to print: the reading experience. Indeed, there is evidence to support the belief that people's previous experiences with linguistic and non-linguistic input play a central role and "strongly shape" their online interpretations of ambiguity in sentences (MacDonald and Hsiao, 2018, p. 176; Alseraye, 2022).

Therefore, we hope that incorporating these L2 Arabic learners as a target population in this continuous research will be helpful in determining the contribution of the two essential explanatory paradigms in sentence parsing/comprehension. This includes the innate explanation (morphological knowledge and word patterns) and the experiential explanation (reading exposure).

The question, then, is, "to what extent the learners of Arabic as a second language, in their advanced competency level, are affected by the incomplete representation of speech (the absence of short vowels and diacritics) in processing GP sentences, their reading time and comprehension." The response should help to uncover the characteristics of the Arabic parser, by determining to what degree the experience factor is an essential variable by itself or in collaboration with the innate variable that plays a major role in reading ambiguous sentences such as the GP sentences. Therefore, the current study targeted Arabic learners with the justification that their reading experiences with Arabic are evolving, and that this should shed some light on the role of experience in reading Arabic ambiguous sentences.

Method

Participants

For the purpose of the study, the sampling technique was judgmental/purposive. A total of 41 participated in the study: 36 of whom were advanced, non-native male Arabic learners of different nationalities, aged from 22 to 26 and enrolled in an Arabic Language Program offered by the Arabic Linguistics Institute at King Saud University, Riyadh, Saudi Arabia. The remaining five participants had already graduated from the Arabic institute.

Initially, to identify participants for the study, three teachers were consulted to assess their students' language competency on a 5 point rating scale (5-excellent, 4-very good, 3-good, 2-fair, 1-poor). For a sample size consideration, only the participants who were in the 1-poor level were excluded right from the start. According to the rating value means, the participants ranged in language competency between 3 and 5, with only 4 out of the 34 participants were rated less than 4. The overall mean was 4.38, with a standard deviation equals 0.652.

The participants were all offered R80 (\$20) as compensation for participating. The data that were later used for the analysis included only those participants who demonstrated a reading fluency skill, based on the post criterion assessment (reading aloud task) that was held after completing the computer task, of whom there were 34. Official approval and consent for participation were obtained beforehand.

Materials

The same two sets of sentences that were constructed as the stimuli for a previous study (Alseraye, 2022) were used for this study and for a subsequent comparison to native Arabic speakers, to assess

assumptions that had been raised by the previous study. The first set of reading conditions included 31 sentences, seven of which were for the practice session, 12 to represent the plain reading condition, rc1, and 12 for the fully vowelized and diacriticized reading condition, rc2. There were four GP sentences in each reading condition. In rc2, when presented as fully vowelized and diacriticized, the short vowels and diacritics on the initial HP-HG words in the sentences would resolve the GP structure only if the readers assembled them with the consonants.

The second set of conditions contained 12 sentences: one for the practice session; the others representing the wrongly vowelized reading condition, rc3. The consonants were supplemented with incorrect short vowels. Four potential GP sentences were included too in this reading condition. For each sentence a textually based comprehension question was constructed that entailed three responses: *true*, *false*, and *I do not know*. The third option, *I do not know*, was given to help the participants avoid having to guess. All the questions were presented fully vowelized and diacriticized.

Only the GP sentences in each condition were the targets, and the remaining stimuli were used as filler items and for a comparison reason.

There were seven words in each sentence. All words were of high frequency, and represent the basic structure Arabic takes, and that Arabic readers encounter in connected texts. The sentences among the three conditions were matched syntactically (see [Appendix B](#)). In the GP sentences, approximately three words separated the initial HP-HG word from the disambiguating region (for an example, see [Diagram 1](#)). The sentences and questions were already assessed and judged by a team of Arabic fourth-grade teachers and graduate students in teaching Arabic program, and then reassessed for the current study by some graduate students in the program of teaching Arabic as a second language. Assessing the sentences and questions was in terms of naturalness, accuracy, suitability, word familiarity, capability of capturing comprehension, and so on. No change in the original sentences was made (see [Appendix A](#) for the sentences and questions used in the experiment).

For a post criterion assessment of reading fluency, and for manifesting what is going on in the L2 Arabic learners' minds as they approach the GP and the potential/resolved GP sentences (by providing them with the right short vowels and diacritics), an informational/expository text of 170 words of high frequency was constructed for the reading aloud task (see [Appendix C](#)). Three GP sentences were inserted in the text. In addition, it included some passive and active HP-HG initials. Two equivalent versions of the same text was constructed. Both versions were the same and presented as plain, except in one version, the HP-HG initials of GP sentences (and the passive sentences) were provided with the right short vowels/diacritics to turn them into a non-GP sentences.

Measures

The following two dependent variables were measured: reading time, measured to the nearest millisecond, and comprehension product, percentage of correct responses. These were measured and coded as *true*, *false*, and *I do not know*. Each correct answer was assigned a 1; all others, false or I do not know, were given a 0.

Procedure

The study followed and adopted the same procedure and paced-reading software that was used in [Alseraye \(2022\)](#) study on Arab children.

The procedure took the following format: the participants logged in, viewed the instructions, and then started the reading task, using a button (space-bar key) that showed every word sequentially when they clicked on it but hid the previous ones. Once they had finished and pressed the space-bar key, a question would pop up with the three options for responding. The same process continued through to the final sentence. When the participants did not know the answer, or felt tempted to guess, they were told to choose "*I do not know*."

They were informed that they would read sentences in which the words were presented with the wrong short vowels, and that assembling the wrong short vowels would lead to constructing words that had no meaning in Arabic; that is, the graphemic form (consonants) of the words was intact, but the phonological aspect was distorted. Assembling only the consonants and ignoring the short vowel signs would result in participants reading a real word in Arabic.

Once the computer-based task was completed, a short reading task was held immediately: a running record by- the researcher was applied while the participants were asked to read aloud a short text that included two GP sentences, similar to the ones conducted in the computer task (and passive sentences). The aim of the task was for the participants to manifest their reading behavior once they encounter an ambiguous GP structure. Further, the task serves to determine whether L2 Arabic learners were conscious of the ambiguity of the GP sentences. The reading aloud texts were randomly assigned to the participants: some participants read the completely plain version while the others read the one with the resolved GP sentences.

It is worth mentioning here that, the task was used to further ensure that the selected participants were indeed at an advanced proficiency level.

Design and analysis

An empirical study with a one-factor within-subjects design was employed to evaluate the effects of the GP structure on it own and in conjunction with short vowels and diacritics on the reading processes of learners of Arabic. Four analyses were conducted, and two separate statistical procedures were employed, the dependent samples *t*-test and the one-way repeated measures analysis of variance. The two tests' assumptions were checked prior to the analyses (i.e., the level of measurement, normality, homogeneity, outliers, sphericity).

Results

Descriptive part

From a descriptive perspective, observing the reading behavior of the Arabic learners revealed very similar patterns that were observed in previous studies conducted on Arabic adults ([Seraye, 2004](#)) and children ([Alseraye, 2022](#)). However, there was one observation that was unique to the L2 Arabic learners. They would apply what they had already experienced; that is, their background knowledge of the verb form they had just bypassed, to the next verb form they encountered

TABLE 1 Results of the *t*-test on the reading times of GP and non-GP sentences.

Non-GP sentences		GP sentences		<i>t</i>	<i>df</i>	<i>p</i>
<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
6,308.6	1,850.6	6,635.6	1,952.1	−2.149	33	0.039

GP, garden path.

but would subsequently figure out that the verb was in a passive form. This experience was then applied to the next passive verb, which they would get right.

Statistical part

The plain GP/non-GP reading condition subdata

To respond to the concerns raised by a previous study (Alseraye, 2022), three types of analyses were conducted on three subsets of data. In the first analysis, GP and non-GP sentences in the plain representation were compared using a dependent samples *t*-test, to detect whether those who were less experienced with print were affected by the GP structure of the sentences. This involved comparing GP and non-GP sentences in terms of reading time and percentage of correct responses.

Reading times analysis

Regarding the data for reading times, the analysis (Table 1) revealed that the GP structure had a significant effect [$t(34) = -2.15$, $p = 0.039$]. The difference in the mean values was roughly 327 ms, which means that it took the participants longer to read the GP sentences than the non-GP sentences ($M = 6,636$ ms for the GP sentences; $M = 6,309$ ms for the non-GP sentences).

Since the analysis included the reading times of both, the correct and incorrect answers data, a subset data of only the correct responses was considered by excluding the incorrect answers from the analysis, in order to have a robust results, using the dependent samples *t*-test. The analysis revealed the same results; a significant effect of the GP sentences (6787.688 ms for the GP sentences vs. 6206.397 ms for the non-GP sentences) on the reading times of the participants [$t(33) = -2.63$, $p = 0.013$]. Due to the existence of outliers, a non-parametric test was conducted, and the same results were revealed ($z = -2.881$, $p = 0.004$).

This result is consistent with the findings of previous studies on Arabic native speakers, both adults (Seraye, 2004) and children (Alseraye, 2022), which demonstrated the effect of the garden-path structure on the reading processes for Arabic texts (see Table 2):

Based on the overall means for the L2 Arabic learners, Arab children, and Arab adults, we found that the GP structure influenced the reading processes of Arabic readers, regardless of their reading levels, reading experience, and print exposure. Another observation concerned the total time spent reading the GP sentences; the Arab adults and the L2 Arabic learners took roughly the same amounts of time to read these sentences (i.e., the same trend). Note, however, that they read different sets of sentences in terms of length and the distance between the HP-HG initial of the GP sentences and their ambiguating region. According to the literature, the decrease in distance between the initial word of the sentence and its disambiguating region should

TABLE 2 Results of the *t*-tests on the reading times of GP and non-GP sentences between the three populations.

Population		Non-GP sentences		GP sentences		<i>p</i>
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
L2 Arabic	Learners	6,308.6	1,850.6	6,635.6	1,952.1	0.039
Arab	Children	8,112.7	2,380.6	8,172.3	2,486.3	0.710
Arab	Adults	6,259.3	1,413.3	6,747.1	2,071.9	0.016

TABLE 3 Results of the *t*-test on the reading comprehension of GP and non-GP sentences.

Non-GP sentences		GP sentences		<i>t</i>	<i>df</i>	<i>p</i>
<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
0.86	0.17	0.88	0.23	−0.362	33	0.720

GP, garden path.

positively affect the reanalysis of the GP sentences (Ferreira and Henderson, 1998; Ferreira et al., 2001), and help to keep the essential parts of the sentence active.

Reading comprehension analysis

In terms of the data for the reading comprehension, the analysis did not reveal any significant differences between the means for the percentages of correct responses for either type of sentences [$t(33) = -0.362$, $p = 0.720$] (see Table 3). The correct responses of the participants did not, on average, differ significantly between the GP and non-GP sentences (the overall mean for the non-GP sentences was $M = 0.86$; the overall mean for the GP sentences was $M = 0.88$).

However, because the data on comprehension were extremely skewed, since the participants' comprehension was generally very good and because of the outliers, a non-parametric test, the Wilcoxon matched-pair signed-rank test, was used along with the dependent samples *t*-test analysis. However, the analysis did not reveal any significant differences between the two means, *z* values (−1.023), and *p*-values (0.306); therefore, only the result of the *t*-test is provided in Table 4.

These results, which show no significant differences between the two types of sentences, are consistent with those of Seraye's two studies of highly skilled adult readers and children who are beginning readers (2004) and (Alseraye, 2022) respectively). Although the GP structure affected reading processing by adding more time loads onto the process, the readers' comprehension was not affected: they had higher scores regardless of the type of structure they processed (Table 4).

The same pattern that emerged from the data for the reading times on the three populations is observed with the comprehension data. The adults, both Arabs and non-Arabs, had higher scores on average on the GP sentences than the Arab children. In addition, the difference between the two means was less among L2 Arabic learners, with only a 2% difference. However, among the Arab adults and Arab children the difference was nearly 7%.

Descriptively, when aligning the comprehension data with the reading time data, a pattern emerges that suggests a relationship between the two (Tables 3, 4).

TABLE 4 Results of the t-tests on the reading comprehension of GP and non-GP sentences between the three populations.

Population		Non-GP sentences		GP sentences		<i>p</i>
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
<i>L2 Arabic</i>	<i>Learners</i>	0.86	0.17	0.88	0.23	0.720
<i>Arab</i>	<i>Children</i>	0.73	0.22	0.80	0.32	0.105
<i>Arab</i>	<i>Adults</i>	0.83	0.08	0.89	0.17	0.053

TABLE 5 Overall means on reading time for GP sentences.

Group	Reading condition (sentence stimuli)	GP sentences	
		<i>M</i>	<i>SD</i>
	Plain (no short vowels or diacritics)	6514.55	1986.81
	Fully vowelized and diacritized	6640.84	1887.05
	Wrong short vowels	7446.72	2430.67

The GP three reading conditions data

The comprehension issues with the GP sentences were addressed by adding the correct short vowels and diacritics to the initial HP-HG words. To further explore this, a one-way repeated measures analysis of variance was conducted specifically on a subset of the data. This aimed to answer whether the inclusion of short vowels and diacritics would significantly affect the reading process of L2 Arabic learners.

For the GP and potential GP sentences, we compared reading times and comprehension across three different reading conditions, including a control condition with incorrect vowelization. The control was included to determine if the addition of incorrect short vowels would impact reading, suggesting that L2 Arabic learners may not rely on sub- and superscript processing, but rather on consonant processing (Seraye, 2004). The central question was whether the addition of short vowels and diacritics would be beneficial in the reading process of GP sentences by L2 Arabic learners.

Reading times analysis

In terms of the data for the reading times, the assumption of sphericity (using Mauchly's Test of Sphericity) was examined first and found to be significant. Therefore, the condition of sphericity was not met, and a nonparametric test was used in addition to the parametric one.

By first using the parametric test, the analysis on the data for reading times revealed a significant difference between the reading conditions ($F(2, 60) = 8.293, p = 0.003$). It took the participants on average 6,514.55 ms to read the GP sentences that were presented as plain ones, 6,640.84 ms to read the GP sentences that were supplemented with the correct short vowels and diacritics, and 7,446.72 ms to read the GP sentences that were supplemented with incorrect short vowels (Table 5 and Figure 1).

Furthermore, a non-parametric test (Friedman Test) showed a significant difference between the mean ranks, $\chi^2(2)$ (Chi-Square) = 6.258, $p = 0.044$.

However, the overall means showed that it took the participants less time on average to read the plain sentences than their

vowelized-diacritized counterparts (126 ms difference). These means also showed that participants took more time on average to read the vowelized condition that was incorrect (see Table 5).

For the same reason mentioned above, the incorrect answer data were excluded from the analysis, using a one-way repeated measures analysis of variance. The analysis of the reading times for the GP sentences revealed the same results; a significant effect of the GP sentences between the reading conditions [$F(2, 60) = 3.749, p = 0.044$]. Due to the violation of the test assumption (Mauchly's Test of Sphericity), a nonparametric test (Friedman Test) was conducted, revealing nonsignificant difference between the mean ranks, $\chi^2(2)$ (Chi-Square) = 5.871, $p = 0.053$. Examining the total means visually shows, however, that the participants took on average less time to read rc1 ($M = 6,635.97$; $SD = 1980.18$) than rc2 ($M = 6,687.60$; $SD = 2075.70$), and much less than rc3 ($M = 7,429.26$; $SD = 2529.75$). Indeed, although it was not specifically an aim in the current study, examining the raw data of only the disambiguating region, word-6 by itself, in the three GP reading conditions and the non-GP plain reading condition showed that the participants, on average, spent longer reading (fixating) the word-6 ($M = 948.60$ ms) relative to its counterparts in the non-GP plain sentences (880.09 ms) and the two reading conditions ($M = 892.25$ m, in rc2; $M = 849.12$, in rc3).

This result is not consistent with Alseraye's (2022) study of Arab children, which showed that they took on average more time to read the plain reading condition and less time to read the vowelized counterparts that were incorrect (Table 6). The data suggest, then, that L2 Arabic learners do not benefit from the short vowels in processing the ambiguous sentences (i.e., GP), but that they were influenced by the effects that the wrong short vowels had on the consonants. As they reported after their sessions reading aloud, this interfered with their reading processes, and they found it difficult to ignore the short vowels and diacritics. This interference was observed among both the native and the non-native participants. As the mean values in Table 6 show, the adults took more time to read the GP sentences that were resolved by the right diacritics.

Reading comprehension analysis

The assumption of sphericity was examined for the reading comprehension data and found not to be significant; the condition of sphericity was met. The repeated measures analysis of variance revealed no significant difference between the three reading conditions [$F(2, 60) = 3.109, p = 0.052$]. However, by accepting that the p value was nearly significant, the pairwise comparisons showed that the only significant difference was between rc1, the plain one, and rc2, the vowelized one ($p = 0.017$). On average, the participants scored best on the plain reading condition ($M = 0.89$), and worst on the vowelized one ($M = 0.77$) (Table 7 and Figure 2).

Since the data were not normally distributed, a non-parametric test was run (Friedman Test), revealing a significant difference between the mean ranks, $\chi^2(2)$ (Chi-Square) = 6.077, $p = 0.048$.

This result, as shown in the raw means in Tables 7, 8, is not completely consistent with the previous studies. The only consistency observed is that the readers of Arabic, native and non-native alike, scored higher when the GP sentences were presented only in consonant form (plain rc1), and that adding the resolving short vowels and diacritics to the consonants of the GP sentences showed no benefits.

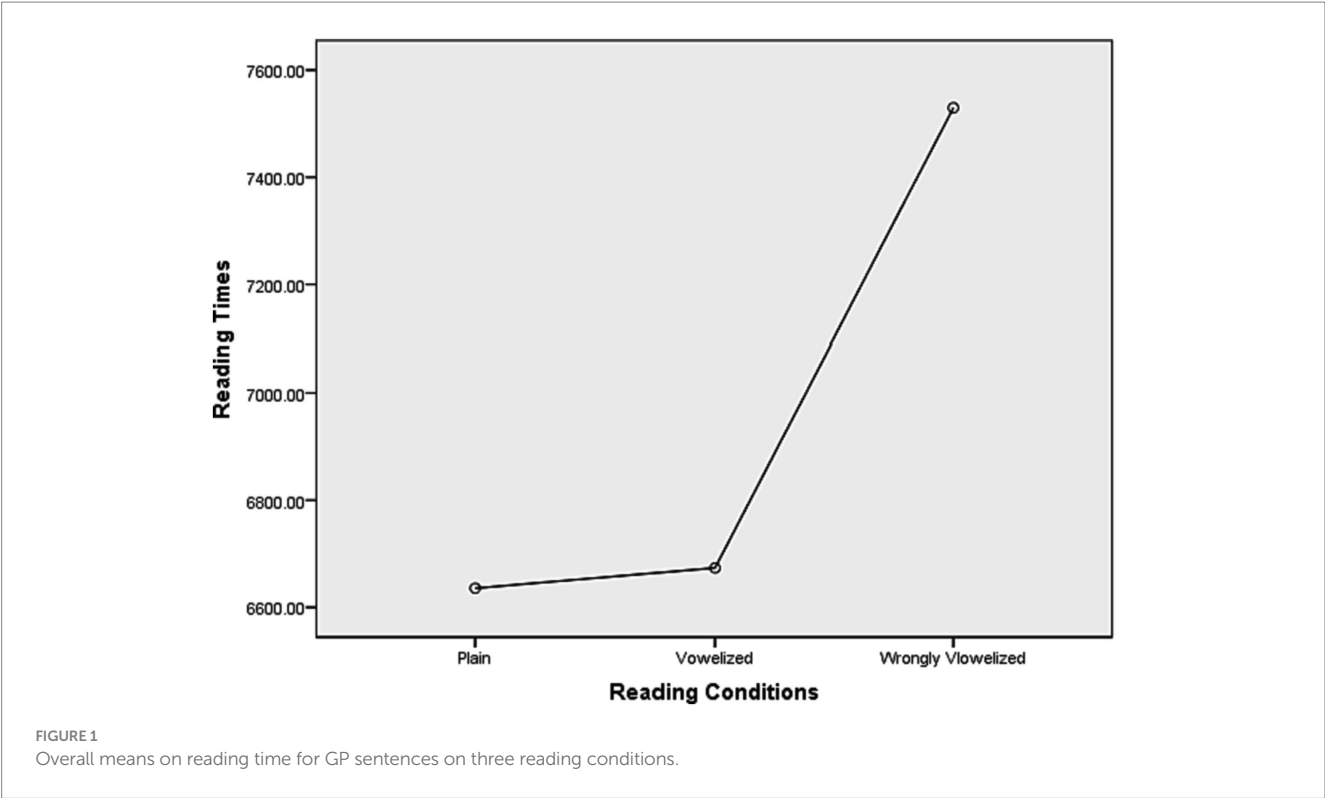


TABLE 6 Results of the repeated measures analysis of variance on the reading times of GP sentences between the three populations.

Population		Plain condition	Vowelized condition	Wrongly vowelized p-value condition	
		M	M	M	
Arabic	Learners (current study)	6514.6	6640.8	7446.7	p = 0.044
Arab	Children (Alseraye, 2022)	8172.3	8007.6	7882.7	p = 0.565
Arab	Adults (Seraye, 2004)	6747.1	6997.3		
			7230.6*		

*These reading time values were for reading conditions in which the GP sentences were provided with only the right diacritics, either a sukun or a case ending (i.e., not vowelized completely) that when assembled would resolve the ambiguity by turning the GP sentence into a non-GP counterpart.

TABLE 7 Overall means on reading comprehension for GP sentences.

	Reading condition (sentence stimuli)	GP sentences	
		M	SD
Group	Plain (no short vowels or diacritics)	0.89	0.21
	Short vowels-plus-diacritics	0.77	0.24
	Wrong short vowels	0.80	0.21

Discussion

Essentially, there are two general observations that summarize the data for reading times and reading comprehension in a consistent manner (Tables 4, 6, 8). The first is that the readers of Arabic spent less time on average reading GP sentences in plain reading conditions, where only the consonants were presented, and that providing them with short vowels and diacritics seems to introduce a disturbing factor by increasing their reading times. This result was reached through

visual examination by Seraye’s (2004) study of highly skilled Arab adults that revealed the following: “the more the short vowels and shaddah signs were provided, the more time it took the participants to read the sentences” (p. 181). This effect that the GP structure had on the reading process, as reflected in the extra time required to read this type of structure, is well documented across different orthographies and various writing systems. It can be explained according to two assumptions: “on the basis of the implicit checking process that operates with a delay cost or on the basis of the processing load in the ambiguous region.” The effects were demonstrated by several other studies that used different techniques such as eye-tracking studies (Roman et al., 1985; Ferreira and Henderson, 1990, Experiment 1; Hermena et al., 2015), first fixation data (Frazier and Rayner, 1982), self-paced reading tasks (Mitchell et al., 1992, Experiment 1, as cited in Mitchell, 1994, p. 381; Seraye, 2004, Experiment 1; Alseraye, 2022), and brain imaging (Mason et al., 2003). The second general observation is that the readers of Arabic scored higher on average on the plain reading condition than for the other conditions (Tables 6, 8). Although providing the consonants

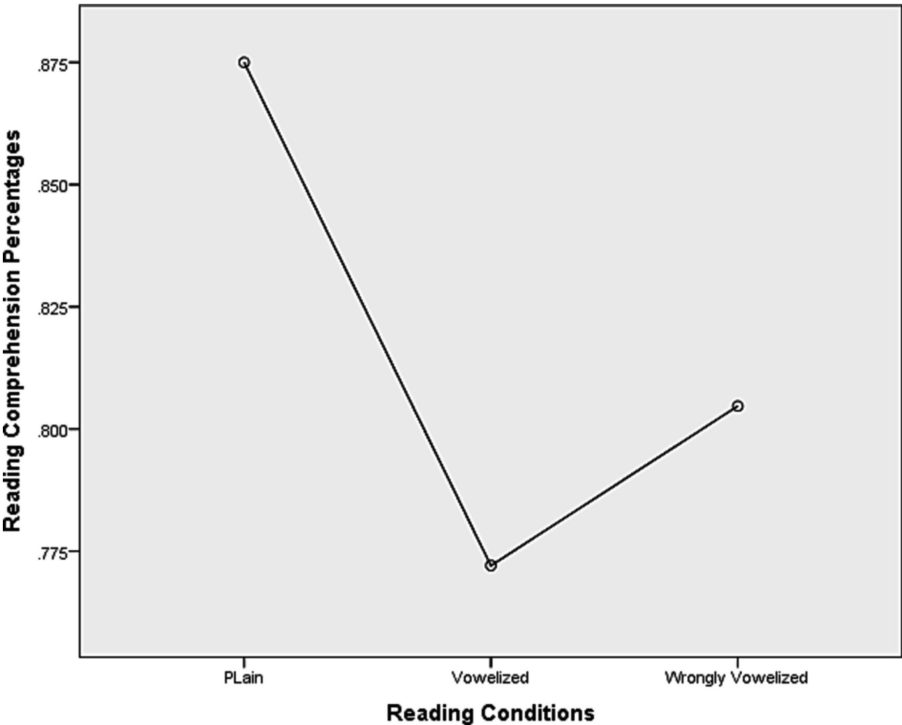


FIGURE 2
Overall means on reading comprehension for GP sentences on three reading conditions.

TABLE 8 Results of the repeated measures analysis of variance on the reading comprehension of GP among the three populations.

Population		Plain condition	Vowelized condition	Wrongly vowelized p-value condition	
		<i>M</i>	<i>M</i>	<i>M</i>	
<i>L2 Arabic</i>	<i>Learners (current study)</i>	0.86	0.73	0.83	<i>p</i> = 0.048
<i>Arab</i>	<i>Children (Seraye, 2022)</i>	0.80	0.69	0.66	<i>p</i> = 0.026
<i>Arab</i>	<i>Adults (Seraye, 2004)</i>	0.89	0.88*	.	
			86		

*These reading comprehension values in the Arabic adult study were for a reading condition in which the GP sentences were provided with only the right diacritics, either a sukun or a case ending (i.e., not vowelized completely), that when assembled would resolve the ambiguity by turning the GP sentence into a non-GP counterpart.

with the right short vowels and diacritics should at least block the GP phenomenon and therefore decrease the reading time of the GP sentences by reducing the hesitancy/reluctance over the disambiguating region, this was not the case. The question then arose regarding the mechanism that helped both the novice and experienced readers to understand GP sentences that were run in a self-paced reading software program design that prevented them from returning to earlier parts of the sentences to clarify or verify their understanding of the GP sentences. One recurrent explanation attributes the good performance to the fact that the readers rely on the richness of Arabic morphology and the pattern, form, and roots in which the words are constituted on either three- or four-root skeletons.

Arabic morphology, which is centered around a trilateral/quadrilateral root system, suggests that Arab readers, when presented with a consonant-based script, are expected to utilize their knowledge of Arabic word formation in accessing mental lexicon representations

(Abu-Rabia, 1995–2001). Seraye (2004) elaborates that within the array of activated potential word forms, there is often a shared trilateral/quadrilateral root indicating a central semantic element, while the word form or pattern (its skeletal tier/word pattern/binyan, McCarthy, 1979, 1981) restricts the potential readings of the word. This structural predictability and the productivity of word forms and affixation are what compensate for the lack of short vowels and diacritics in the written language.

The role of Arabic morphology roots in the reading process has recently, as noted earlier, been revisited, consolidated, and thoroughly documented (Alseraye, 2022, p. 17). Indeed, the investigation was recently directed toward concerns about the Arabic lexicon representation, the classes of morphological representation (roots vs. words patterns: nominal and verbal), and the degree to which the process of naming words could be facilitated (see, for example, Aljasser, 2020; Khateb et al., 2022; Aldholmi and Pycha, 2023).

However, the insignificant role of short vowels and diacritics in the reading processing of GP sentences can also be explained according to two assumptions:

“by the fact that subjects, as [Ferreira et al. \(2009\)](#) state, ‘have a tendency to sacrifice reanalysis of the garden-path in order to keep up with later material. This pattern of results is consistent with the assumptions of the good enough theory of language processing, which assumes that processing resources are limited, and therefore predicts that garden-path reanalysis processes will be curtailed if upcoming material must also be processed’” (p. 416).

As noted earlier, giving up the reanalysis of the GP structure was observed visually during the task involving reading aloud. The participants, Arab adults, Arab children, and L2 Arabic learners, did not go back to reanalyze the GP sentences even when they knew that their initial interpretations of the GP sentences were wrong, although some responded to their mistakes by making exclamations such as, “لا!” meaning “no!” However, the claim that there is no reanalysis was observed was based on the fact that the participants, Arabs and none Arabs, both children and adults, never went back to reread the HP-HG verb, and choose its right form. Indeed, the reanalysis could have been occurred with no trace of verbalizing it, which cannot be examined by using a reading aloud task. Only with an eye-movement technique, such a claim can be assuredly assessed.

The second assumption involves previous reading experience, which is further implicated in the equivocal results ([Alseraye, 2022](#)). Monitoring the reading behavior of the Arab and non-Arab readers of the language during the task involving reading aloud demonstrates that they apply what they have already experienced to the next verb form they meet in the text.

Another explanation that can be presented here, and which is supported by the data of both the current study and previous studies by [Seraye \(2004\)](#) and [Alseraye \(2022\)](#) is related to predictions. That is, because of the segregability representation of Arabic writing system, and the absence of short vowels and diacritics, the readers would expect to be able to predict what is next, based on the semantics and syntax of the language during reading an HP-HG initial sentence. In the Arabic case, the readers are expected to use their knowledge of the semantic and syntactic features in predicting what follows. Within the sentence parsing models (Left-Corner parsing, the Garden-Path model, Syntactic Prediction Locality Theory (SPLT), Good-Enough and Noisy Channel processing, and Surprisal and Entropy in Information-Theoretic models of language processing), prediction is considered a central component in modeling human sentence parsing ([Ferreira and Qiu, 2021](#)). Although both semantics and syntax are clear predictors in explaining the situation, the syntactic prediction in the case of Arabic seems to be strong logically and through observation because of the different characteristics of the language, including its morphological features. These include the dominant word order, the features of the writing style (anastrophe), and the inversion of the word orders in phrases and clauses resulting from the disappearance of grammatical case endings, which would force the writer to avoid any disturbance that could arise through this disappearance. Furthermore, in any modern Arabic writing, the distance between the subject and its predicate is not too far to put

a load on the reader’s memory. To illustrate, the transitive forms of the verbs are sometimes used as intransitive forms, and prepositions are attached to the NPs to help the readers to grasp the focus in advance and avoid any associated disturbance. Arabic writing includes many transitive verbs that are currently used as intransitive ones. For example, there is the verb “قَبِلَ” meaning “[He] accepted ...,” in which the transitive verb has become intransitive in journalistic writing [for more details on this issue, see [Afifi Ahmed’s \(2004\)](#) study].

Also, with the passive voice for verbs, although Arabic allows both aspects, the active and passive voices, it is expected that the stylistic features of the discourse help in finding the aspect voice of the verb. Indeed, even having the verb next to its subject would be close enough to prevent the ambiguity, particularly with respect to the GP phenomenon. However, using the passive voice, where the GP phenomenon would be obvious, is not really encouraged unless the context requires it. Furthermore, in modern Arabic writing, alien/outlandish expressions can leak into the writers’ linguistic expressions when they want to avoid using the passive voice by inserting the word “تَمَّ” which means “was completed,” which acts as a suffix, and converting the main verb “كسر” into a gerund. This, then, means that the object “الكرسي” the chair” was acted upon by some other performer of the verb, or that they may use other morphological devices by changing the patterns of the verbs, while keeping the meaning unchanged, such as through the form pattern “انكسر” broke itself” or using the topicalization mechanism (see the examples below):

خالد كسر الكرسي Khalid broke the chair.
 الكرسي كُسر بواسطة خالد/ كُسر الكرسي بواسطة خالد The chair was broken by Khalid.
 تم كسر الكرسي The chair has been broken.
 انكسر الكرسي The chair was broken.

Clearly, this type of writing representation would affect readers, and they would, therefore, expect to follow a trend of reading behavior. The extended exposure to the current print as it is usually represented would be likely to build a determined experience in its readers that would eventually help them to construct in their minds a faculty of prediction, which would then show in their reading behaviors. This appears reasonable because of the incompleteness of speech in Arabic print that is due to the absence of short vowels and diacritics from the print. Using this type of experience would help prepare the reader to emphasize some sensory inputs and ignore others. That is, the cognition of the experienced readers of Arabic, to apply Gibson’s theoretical framework of perception, would be built on a foundation of perceptual knowledge that would become a system of representation about the verbal sentences that begins with an HP-HG initial word ([Gibson, 1988](#)). Indeed, the constructivist view of perception presented by [Gregory \(1997\)](#) might be a good grounding explanation for what the Arab readers do while processing HP-HG initial sentences. That is, a top-down perceptual view is emphasized in which the visual perception of Arabic readers uses inferences from visual cues and past experience during processing HP-HG initial sentences.

To illustrate this, we observed that both native and L2 Arabic learners read the initial HP-HG verbal word of a sentence as a verb in the active voice (the default in their minds), even when the initial verbal word was supplied with the necessary short vowels and diacritics. Therefore, because speech is less represented in Arabic orthography, the parser would be trained to rely on past experiences (e.g., frequent exposures) and approach those homographic-initial word sentences first.

Based on the findings of this study, it is recommended that different techniques to be used on the same sentence stimuli, to consolidate or refute the current findings as, for example, by adopting the eye movement technique. Other recommendations are the adoption of a qualitative approach in figuring out what is going on in the Arabic readers' mind, as they approach and resolve the ambiguity caused by the GP sentences, in their two types of representations: vowelized and non-vowelized text representations. That is, by using the "Think Aloud" procedure, the participants should be asked to verbalize their thinking processes as they read GP sentences. It is also recommended that different populations to be targeted; for example, by investigating the effect of the GP phenomenon on less skilled readers and on participants who have reading difficulties, such as dyslexia.

Pedagogically, the ambiguity resulting from the omission of short vowels and diacritics should be addressed in Arabic teaching contexts. Given their ubiquitous absence in printed material, it is beneficial for Arabic learners to practice reading texts, particularly the HP-HG initial sentences, with such orthographic characteristics. Furthermore, students should be taught to apply short vowels and diacritics both economically and efficiently in their own writing.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Research Ethics Committee-Vice Rectorate for Graduate Studies & Scientific Research-Deanship of Scientific Research-King Saud University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

References

- Abu-Rabia, S. (1995). Learning to read in Arabic: reading, syntactic, orthographic & working memory skills in normally achieving and poor Arabic readers. *Read Psychol.* 16, 351–394. doi: 10.1080/0270271950160401
- Abu-Rabia, S. (2012). The role of morphology and short vowelization in reading morphological complex words: evidence for the domination of the morpheme/root-based theory in Reading Arabic. *Creat. Educ.* 3, 486–494. doi: 10.4236/ce.2012.34074
- Abu Rabia, S. (2019). The role of short vowels in reading Arabic: a critical literature review. *J. Psycholinguist. Res.* 48, 785–795. doi: 10.1007/s10936-019-09631-4
- Abu-Rabia, S. (2021). The Impact of Diacritic Marking on Listening Comprehension Skills in Arabic Orthography. *The Journal of Educational Research.* 114, 64–73. doi: 10.1080/00220671.2021.1872476
- Abu-Rabia, S., and Abu-Rahmoun, N. (2012). The role of phonology and morphology in the development of basic Reading skills of dyslexic and Normal native Arabic readers. *Creat. Educ.* 3, 1259–1268. doi: 10.4236/ce.2012.37185
- Afifi Ahmed, A. (2004). *يُرى صِلًا فاحصلًا يف موزلو لعفلا يدعت يف تارظن* (II) Necmettin Erbakan Üniversitesi İlahiyat Fakültesi Dergisi, sayı: 38, 2004, ss. 101–140.
- Aldholmi, Y., and Pycha, A. (2023). Segmental contributions to word recognition in Arabic sentences. *Poznan Stud. Contemp. Linguist.* 59, 257–287. doi: 10.1515/pscl-2022-2010
- Al-Fahid, J. M. (2000). *The Goodman Psycholinguistic Model of English Reading and Its Applicability to Semitic languages*. Unpublished Doctoral Dissertation, University of Arizona.
- Aljasser, F. (2020). Root and pattern effects in the processing of spoken non-words in Arabic. *Int. J. Linguist.* 12:292. doi: 10.5296/ijl.v12i2.16545
- Alseraye, A. M. (2022). The Reading comprehension process of Arabic garden-path sentences: an equivocal effect. *Creat. Educ.* 13, 3026–3052. doi: 10.4236/ce.2022.139191
- Boudelaa, S., and Marslen-Wilson, W. (2015). Structure, form, and meaning in the mental lexicon: evidence from Arabic. *Lang. Cognit. Neurosci.* 30, 955–992. doi: 10.1080/23273798.2015.1048258

Author contributions

AA: Data curation, Formal analysis, Methodology, Supervision, Writing – original draft, Investigation, Project administration, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

I would like to express my thanks and deepest gratitude to the Arabic Linguistics Institute in King Saud University, their administration and students, for the assistance and warm welcome I was given.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2024.1333112/full#supplementary-material>

- De Francis, J. (1989). *Visible speech: The diverse oneness of writing systems*. Honolulu: University of Hawaii
- El Akiki, C., and Content, A. (2020). Early sensitivity to morphology in beginning readers of Arabic. *Front. Psychol.* 11:552315. doi: 10.3389/fpsyg.2020.552315
- Ferreira, F., Christianson, K., and Hollingworth, A. (2001). Misinterpretations of garden-path sentences: implications for models of sentence processing and reanalysis. *J. Psycholinguist. Res.* 30, 3–20. doi: 10.1023/A:1005290706460
- Ferreira, F., Engelhardt, P. E., and Jones, M. W. (2009). “Good enough language processing: a satisfying approach” in *Proceedings of the 31st annual conference of the cognitive science society*. eds. N. Taatgen, H. Rijn, J. Nerbonne and L. Schomaker (Austin, TX: Cognitive Science Society), 413–418.
- Ferreira, F., and Henderson, J. M. (1990). Use of verb information in syntactic parsing: evidence from eye movements and word-by-word self-paced reading. *J. Exp. Psychol. Learn. Mem. Cogn.* 16, 555–568.
- Ferreira, F., and Henderson, J. M. (1998). “Syntactic reanalysis, thematic processing, and sentence comprehension,” in *Reanalysis in sentence processing*. eds. J. D. Fodor and F. Ferreira (Dordrecht: Kluwer), 73–100.
- Ferreira, F., and Qiu, Z. (2021). Predicting syntactic structure. *Brain Res.* 1770:147632. doi: 10.1016/j.brainres.2021.147632
- Frazier, L., and Rayner, K. (1982). Making and correcting errors in sentence comprehension: eye movements in the analysis of structurally ambiguous sentences. *Cogn. Psychol.* 14, 178–210. doi: 10.1016/0010-0285(82)90008-1
- Gibson, E. J. (1988). Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. *Annu. Rev. Psychol.* 39, 1–42. doi: 10.1146/annurev.ps.39.020188.000245
- Gregory, R. L. (1997). Knowledge in perception and illusion. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 352, 1121–1127. doi: 10.1098/rstb.1997.0095
- Hermena, E., and Reichle, E. (2020). Insights from the study of Arabic Reading. *Lang. Linguist. Compass* 14, 1–26. doi: 10.1111/lnc3.12400
- Hermena, E. W., Drieghe, D., Hellmuth, S., and Liversedge, S. P. (2015). Processing of Arabic diacritical marks: phonological-syntactic disambiguation of homographic verbs and visual crowding effects. *J. Exp. Psychol. Hum. Percept. Perform.* 41, 494–507. doi: 10.1037/xhp0000032
- Ibrahim, R. (2013). Reading in Arabic: new evidence for the role of vowels diacritics. *Creat. Educ.* 4, 248–253. doi: 10.4236/ce.2013.44036
- Khateb, A., Asadi, I. A., Habashi, S., and Korinth, S. P. (2022). Role of morphology in visual word recognition: a Parafoveal preview study in Arabic using eye-tracking. *Theory Pract. Lang. Stud.* 12, 1030–1038. doi: 10.17507/tpls.1206.02
- MacDonald, M. C., and Hsiao, Y. L. (2018). “Sentence Comprehension,” in *The Oxford Handbook of Psycholinguistics*. eds. S.-A. Rueschemeyer and M. G. Gaskell (Oxford University Press), 171–196.
- MacDonald, M., Pearlmutter, N., and Seidenberg, M. (1994). The lexical nature of syntactic ambiguity resolution. *Psychol. Rev.* 101, 676–703. doi: 10.1037/0033-295X.101.4.676
- Mahfoudhi, A. (2007). “Roots and patterns in Arabic lexical processing” in *Perspectives on Arabic linguistics XIX*. ed. E. Benmamoun (Amsterdam: John Benjamins), 123–150.
- Mahfoudhi, A., Elbeheri, G., Al-Rashidi, M., and Everatt, J. (2010). The role of morphological awareness in reading comprehension among typical and learning disabled native Arabic speakers. *J. Learn. Disabil.* 43, 500–514. doi: 10.1177/0022219409355478
- Marcus, G. F., Brinkmann, U., Clahsen, H., Wiese, R., and Pinker, S. (1995). German inflection: The exception that proves the rule. *Cognitive Psychology*, 29, 189–256. doi: 10.1006/cogp.1995.1015
- Maroun, M. (2017). Diacritics and the resolution of ambiguity in Reading Arabic. doctoral dissertation, University of Essex.
- Mason, R., Just, M., Keller, T., and Carpenter, P. (2003). Ambiguity in the brain: what brain imaging reveals about the processing of syntactically ambiguous sentences. *J. Exp. Psychol. Learn. Mem. Cogn.* 29, 1319–1338. doi: 10.1037/0278-7393.29.6.1319
- McCarthy, J. (1979). Formal problems in Semitic phonology and morphology. Doctoral dissertation, MIT, Cambridge, Massachusetts.
- McCarthy, J. (1981). A prosodic theory of nonconcatenative morphology. *Linguist. Inquiry* 12, 373–418.
- Mitchell, D. C. (1994). “Sentence parsing” in *Handbook of psycholinguistics*. ed. M. A. Gernsbacher (San Diego: Academic Press), 375–409.
- Mitchell, D. C., Corley, M. M. B., and Garnham, A. (1992). Effects of context in human sentence parsing: Evidence against a discourse-based proposal mechanism. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 69–88. doi: 10.1037/0278-7393.18.1.69
- Mohammad, M. (2000). *Word order, agreement in Arabic*. Amsterdam: John Benjamins B. V.
- Ots, N. (2021). Cognitive constraints on advance planning of sentence intonation. *PLoS One* 16:e0259343. doi: 10.1371/journal.pone.0259343
- Roman, G., Pavard, B., and Asselah, B. (1985). Traitement perceptif des phrases ambiguës en Arabe [perceptive analysis of ambiguous Arabic sentences]. *Cahiers Psychol Cognit* 5, 5–22.
- Saiegh-Haddad, E. (2017). “Learning to read Arabic” in *Learning to read across languages and writing systems*. eds. L. Verhoeven and C. Perfetti (Cambridge: Cambridge University Press), 104–126.
- Seraye, A. M. (2004). The role of short vowels and context in the reading of Arabic, comprehension, and word recognition of highly skilled reader. Doctoral dissertation, University of Pittsburgh
- Seraye, A. M. (2016). Short vowels versus word familiarity in the reading comprehension of Arab readers: a revisited issue. *Int. Electron. J. Element. Educ.* 8, 418–506.
- Taha, H. (2016). Deep and shallow in Arabic orthography: new evidence from reading performance of elementary school native Arab readers. *Writ. Syst. Res.* 8, 133–142. doi: 10.1080/17586801.2015.1114910
- Taha, H., and Saiegh-Haddad, E. (2016). The role of phonological versus morphological skills in the development of Arabic spelling: an intervention study. *J. Psycholinguist. Res.* 45, 507–535. doi: 10.1007/s10936-015-9362-6
- Taha, H., and Saiegh-Haddad, E. (2017). Morphology and spelling in Arabic: development and interface. *J. Psycholinguist. Res.* 46, 27–38. doi: 10.1007/s10936-016-9425-3
- Wattad, H., and Abu-Rabia, S. (2020). The advantage of morphological awareness among normal and dyslexic native Arabic readers: a literature review. *Read. Psychol.* 41, 130–156. doi: 10.1080/02702711.2020



OPEN ACCESS

APPROVED BY
Frontiers Editorial Office,
Frontiers Media SA, Switzerland

*CORRESPONDENCE
Abdullah M. Seraye Alseraye
✉ aalseraye@ksu.edu.sa

RECEIVED 08 April 2024
ACCEPTED 10 April 2024
PUBLISHED 22 April 2024

CITATION
Alseraye AMS (2024) Corrigendum: L2 Arabic
learners' processing of Arabic garden-path
sentences: a consistent reading pattern.
Front. Psychol. 15:1414363.
doi: 10.3389/fpsyg.2024.1414363

COPYRIGHT
© 2024 Alseraye. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License \(CC
BY\)](#). The use, distribution or reproduction in
other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Corrigendum: L2 Arabic learners' processing of Arabic garden-path sentences: a consistent reading pattern

Abdullah M. Seraye Alseraye *

Department of Curriculum & Instruction, School of Education, King Saud University, Riyadh, Saudi Arabia

KEYWORDS

garden-path structure, Arabic short vowels, past experience, heterophonic-homographic initial, L2 Arabic learners, reading comprehension

A corrigendum on

L2 Arabic learners' processing of Arabic garden-path sentences: a consistent reading pattern

by Alseraye, A. M. S. (2024). *Front. Psychol.* 15:1333112. doi: 10.3389/fpsyg.2024.1333112

In the published article, there was an error. The author realized that a mistake was made during the process of writing the article. The author name quoted, "Ots," was repeated in the next sentence, leading to a claim being mistakenly attributed to this author. A correction has been made to clarify the text and amend the mistaken attribution.

A correction has been made to *Introduction, Paragraph 6*. The corrected sentence appears below:

According to Ots (2021), the speaker, in the linguistic encoding stage in language production, would "assign the syntactic functions [that] are appropriate for the message and order the constituents, given the discourse and grammatical constraints" (p. 2). Similarly, the writer, I assume, would be forced to comply with these constraints, in addition to the cognitive constraints and limitations, to avoid the long distance between the HP-HG word and its disambiguating region in the sentence.

The author apologizes for this error and states that this does not change the scientific conclusions of the article in any way. The original article has been updated.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.



OPEN ACCESS

EDITED BY

Oksana Laleko,
State University of New York at New Paltz,
United States

REVIEWED BY

Alina Konradt,
University College London, United Kingdom
Natalia Mitrofanova,
UiT The Arctic University of Norway, Norway
Marina Sokolova,
Norwegian University of Science and
Technology, Norway

*CORRESPONDENCE

Natalia Slioussar
✉ slioussar@gmail.com

RECEIVED 25 November 2023

ACCEPTED 05 March 2024

PUBLISHED 20 March 2024

CITATION

Slioussar N and Harchevnik M (2024) Word
order and context in sentence processing:
evidence from L1 and L2 Russian.
Front. Psychol. 15:1344366.
doi: 10.3389/fpsyg.2024.1344366

COPYRIGHT

© 2024 Slioussar and Harchevnik. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Word order and context in sentence processing: evidence from L1 and L2 Russian

Natalia Slioussar^{1,2*} and Maria Harchevnik²

¹School of Linguistics, National Research University Higher School of Economics, Moscow, Russia,

²Institute for Cognitive Studies, Saint Petersburg State University, Saint Petersburg, Russia

Introduction: In this paper, we studied how native (L1) speakers of Russian and speakers of Mandarin Chinese learning Russian as a foreign language (L2) process Russian sentences with different word orders. We compared SVO (canonical) and OVS (non-canonical) orders in isolation and in context. Experiments focusing on the L2 processing of different word orders are still not very numerous, and those using context are extremely rare.

Methods: In Experiment 1, target sentences were presented in isolation. In Experiment 2, one-sentence contexts introduced one NP mentioned in the target sentence, either the first (so that given information preceded new information, which is characteristic for Russian and many other languages) or the second. As a result, two factors could be compared: the syntactic (word order) and the contextual (whether the context is appropriate from the information-structural perspective). We used different measures to capture online and offline effects: word-by-word reading times, question-answering accuracy and sentence rating on a 1 to 5 scale (for L1 participants).

Results and discussion: In both experiments, RTs and question-answering accuracy data showed that non-canonical orders were difficult for L2 participants, but not for L1 participants. However, L1 participants gave non-canonical orders lower ratings in isolation, presumably because in naturally occurring texts, they are used only in particular contexts. As for the context factor in Experiment 2, some effects were the same for L1 and L2 processing: all participants read given NPs faster than new ones and preferred sentences with a 'given – new' word order. The latter may reflect the universal principles of narrative coherence. However, unlike native speakers, L2 readers are not sensitive to more subtle contextual requirements of different word orders.

KEYWORDS

word order, context, information structure, second language processing, Russian

1 Introduction

In most languages, different word orders are possible in a sentence, although such alternations are more diverse and more widespread in some languages than in the others. They are primarily associated with the information structure of the sentence (which information is new or given, salient or backgrounded). One word order (the most frequent, with the least specific information-structural requirements) is termed canonical or basic. Many studies are dedicated to processing sentences with different word orders, both by native (L1) speakers and by second language (L2) learners, for whom acquiring the rules underlying word order alternations was shown to be particularly difficult in various languages.

In the first processing experiments with L1 readers, different orders were presented in isolation. But, since their use depends on information structure, subsequent studies presented them in contexts. L2 processing experiments are still not very numerous, and those using contexts are extremely rare. However, it would be interesting to find out whether L2 readers are sensitive to various contextual requirements of different orders, and how they differ from L1 readers in this respect.

In the present study, we aimed to fill this gap. We compared how native speakers of Russian and speakers of Mandarin Chinese learning Russian as a foreign language process Russian sentences with different word orders. In Experiment 1, these sentences were presented in isolation, while in Experiment 2, we used one-sentence contexts satisfying or violating information-structural requirements of these sentences. We used different measures (word-by-word reading times, question-answering accuracy and sentence rating on a 1–5 scale) to capture online and offline effects.

The structure of the paper is as follows. In the next section, we briefly introduce the main properties of word order in Russian and in Chinese. Then we give an overview of L1 and L2 processing studies focusing on word order alternations. After that, we turn to the present study.

1.1 Word order in Russian and in Mandarin Chinese

We selected sentences with a subject NP, an object NP and a transitive verb for our study. Both in Russian and in Mandarin Chinese (Putonghua), the basic word order is SVO (subject – verb – object) in such sentences (Dryer, 2005).¹ Both languages allow for certain word order alternations, mainly triggered by the information structure. Both languages, especially in written texts, prefer to put given information before new information, when it is possible to change the word order in the sentence accordingly. However, the two languages are very different in other respects: while Russian is a morphologically rich inflected language with morphological case marking, Chinese is an isolating language: most words consist of a single morpheme and have no inflectional morphology. Russian allows for more diverse word order alternations, and in general, possible word orders in these languages are not similar to each other.

In a Russian sentence with a subject, an object and a verb, all six computationally possible orders are attested. Russian has six cases, and does not use prepositions not only with direct objects, but often also with various indirect ones. Due to morphological case marking, subjects, direct and indirect objects can usually be told apart unambiguously.² Slioussar and Makarchuk (2022) conducted a corpus study showing the prevalence of these six orders in more formal and less formal written and oral texts. The basic SVO order clearly prevails everywhere. The second most frequent order in narrative written texts is OVS, and this was one of the reasons to choose it for our experiments.

Word order alternations in Russian and information-structural requirements associated with them were studied by many authors working in different frameworks (e.g., Sirotinina, 1965; Kovtunova, 1976; Krylova, 1992; Bailyn, 1995; Yanko, 2001; Slioussar, 2007; Titov, 2017, 2020). For some orders, these requirements are easier to formulate, while for the others, they are a matter of debate [for example, Slioussar and Makarchuk (2022) discuss this problem for SOV orders]. OVS orders are relatively transparent in this respect, and this was the second reason to choose them. In the majority of cases, they are used when the subject is in focus (new information), while the object is topicalized (usually given).

Apart from information structure, the choice of word order may be affected by the argument prominence hierarchy: humans > animals > inanimates (Titov, 2017; Vihman and Nelson, 2019). Namely, when arguments have the same information-structural status (e.g., are both new), a non-canonical word order may still be used so that NPs denoting humans could precede NPs denoting animals and inanimate things. Since we wanted to focus on information structure in this study, we balanced arguments for animacy in our target sentences.

If the argument prominence hierarchy is controlled for, the canonical SVO order is the only one that is fully appropriate to use in isolation. The widest range of contexts is associated with it. In particular, when the subject is new and the object is given, Russian speakers can use not only OVS, but also SVO, shifting the main stress on the subject, as in (1)–(2) (the NP bearing the main stress is underlined). However, this is more characteristic for dialogues than for narrative texts (Kodzasov, 1996).

(1) *Kto videl Petju?*

who_{NOM.SG} saw Petya_{ACC.SG}.
'Who saw Petya?'

(2) a. *Petju videl Vasja* (OVS)

Petya_{ACC.SG} saw Vasya_{NOM.SG}.

b. *Vasja videl Petju* (SVO)

Vasya_{NOM.SG} saw Petya_{ACC.SG}.
'Vasya saw Petya.'

There were also other reasons to include the OVS order in our experiments. As we show below, such orders are virtually unattested in Chinese, which creates a challenge for Chinese learners of Russian. In addition to that, we wanted to have sentences with an inverted order of arguments — figuring out the predicate-argument structure appears to be the main problem associated with processing of non-canonical orders. Thus, we had SVO and OVS sentences in our study.

Mandarin Chinese (Putonghua) is an isolating language. Therefore, while Russian can rely on case marking to tell arguments apart, word order and context play a crucial role for this in Chinese. Nevertheless, Chinese allows for certain word order alternations (Sun and Givón, 1985; Sun, 2006). In particular, SOV orders are very widespread, especially in spoken language in northern dialects (Li and Thompson, 1974; Gao, 2008). Sentences in which a non-subject NP precedes the verb, while the subject follows it are possible, but with various intransitive verbs (Gao, 2008): these NPs may denote location, time etc. Transitive verbs selecting direct objects are not used in OVS orders, which may create specific problems with Russian OVS sentences for Chinese L2 learners.

¹ Here and below, we discuss affirmative sentences, questions may be different in this respect.

² Sometimes this is not possible due to case syncretism (for example, accusative forms coincide with nominative ones in most inanimate nouns), and the reader has to rely on context and on the primacy of the canonical word order. We avoided such cases in our study.

1.2 L1 processing of sentences with different word orders

Many studies on different languages focused on L1 processing of different word orders. In most of them, target sentences appeared in isolation (Frazier and Flores d'Arcais, 1989; Hyönä and Hujanen, 1997; Gibson, 1998; Bader and Meng, 1999; Clahsen and Featherston, 1999; Sekerina, 1999; Stojanovic, 1999; Miyamoto and Takahashi, 2000; Vasishth, 2002; Erdocia et al., 2009, among others). Non-canonical orders were found to be more difficult to process than canonical ones, although these differences did not reach significance in some studies — presumably, due to the fact that non-canonical orders may be very widespread in some languages (although they are still much less frequent than canonical ones).

However, non-canonical orders have contextual requirements and sound less natural in isolation. To find out to what extent processing difficulties may be associated with word order alternations *per se* and with context, several authors introduced context sentences in their experiments. Let us look at some of these studies in more detail.

Bornkessel et al. (2003) presented German sentences with different word orders in isolation and using questions as contexts. ERPs and word-by-word reading times were recorded. Non-canonical orders were more difficult than canonical ones in isolation. Context eliminated this difficulty, but only partially. Some signature effects associated with the syntactic and contextual factor were identified. These results were supported and extended in later studies (Schumacher and Hung, 2012; Burmester et al., 2014).

Kaiser and Trueswell (2004) examined Finnish sentences with SVO and OVS orders. Their syntactic and information-structural properties are similar to those in Russian, so this study is especially relevant for us. In the first experiment, Kaiser and Trueswell presented these orders after two-sentence contexts, as in (3a–b), and measured word-by-word reading times. The second context sentence introduced one NP from the target sentence, either the first one, as in (3c) (creating a 'given – new' order in the target sentence), or the second one, as in (3d) (creating a 'new – given' order in the target sentence, not characteristic for narrative texts in Finnish, like in Russian). As a result, they could compare the effects of the word order and context type factors. Both factors were significant: sentences in 'new – given' contexts and sentences with non-canonical word orders were read more slowly.

(3) a. *Lotta etsi eilen sieniä metsässä.*

'Lotta looked for mushrooms yesterday in the forest.'

b. *Hän huomasi heinikossa jäniksen joka liikkui varovasti eteenpäin.*

(s)he noticed grass_{LOC} hare_{ACC} that was-moving carefully forward.

c. *Jänistä seurasi hiiri ja linnut lauloivat.*

hare_{PART} followed mouse_{NOM} and birds were-singing.

d. *Hiiri seurasi jänistä ja linnut lauloivat.*

mouse_{NOM} followed hare_{PART} and birds were-singing.

In the second experiment, participants' eye movements were tracked as they looked at stimulus pictures and listened to their descriptions (including context sentences and target SVO or OVS sentences). If the first NP in the target sentence referred to a given referent (that was mentioned in the preceding context), sentences with the OVS order demonstrated anticipatory eye movements toward the discursively new referent even before the participants received sufficient acoustic information to recognize the second NP. This was

not the case for the SVO condition. This shows that Finnish speakers expect the OVS order to be used in certain contexts, while the contextual requirements of the canonical SVO order are much wider.

Sekerina (2003) was the first to compare Russian sentences with different word orders in isolation and in context. However, in her study, one-sentence contexts which did not vary across conditions. A general facilitative effect of context was reported, but non-canonical orders still had longer reading times than canonical ones.

Slioussar's (2011a) study on Russian followed the same logic as the first experiment by Kaiser and Trueswell (2004): 'given – new' and 'new – given' contexts were used. However, Slioussar compared more complex word orders with three argument NPs (a subject, a direct and an indirect object) and had more complex contexts where two out of three NPs were introduced. The context factor was significant, while the word order factor was not: all orders were equally easy to read in an appropriate context. Having longer sentences, Slioussar also could describe in more detail how different contexts affect processing word-by-word.

In several studies on Spanish, Gattei et al. (2015, 2017, 2021) focused on another aspect of processing different word orders: on the problem of establishing predicate-argument structure. In Spanish, like in Russian (e.g., Slioussar, 2011b), agentive subjects tend to precede patientive objects (resulting in the prevalence of SVO orders with active verbs), but patientive subjects tend to follow experiencer objects (resulting in the prevalence of OVS orders in the relevant group of psych verbs).³ Gattei and colleagues demonstrated that these two groups of verbs have distinct processing patterns. In particular, Gattei et al. (2021) used the same two types of contexts, as Kaiser and Trueswell (2004) and Slioussar (2011a), but compared SVO and OVS orders in these two groups of verbs in an eye-tracking-while-reading study. Several diverse measures were used in the study ('early' and 'late' eye-movement measures, accuracy and response times to comprehension questions), and all three factors played a significant role at least for some of them.

1.3 L2 processing of sentences with different word orders

In this section, we will first discuss some general ideas that may be important for our study and then the experiment by Laleko (2022) that is especially relevant for us. Many authors, especially in formal approaches to second language acquisition, have noted that various phenomena at the interface between the grammar and information structure present a challenge even to advanced L2 learners (e.g., DeKeyser, 2005; Callies, 2009; Sorace, 2011). It is easier to master grammatical rules underlying various constructions than to grasp how these constructions are used depending on the discourse context. Sorace generalizes this insight in her Interface Hypothesis (Sorace and Serratrice, 2009; Sorace, 2011, 2012). She assumes that language processing is modular, so it can be expected that using the information within the computational system is easier than figuring out the interactions between the modules. Moreover, external interfaces (e.g., syntax interacting with discourse) are expected to be more challenging

³ This group is not very numerous, so it does not undermine the overall prevalence of the SVO order.

than internal ones (e.g., lexicon interacting with syntax). Notably, learners' difficulties may often be observed only in online tasks because integrating grammatical and discourse information in real-time processing requires more cognitive resources.

The largest number of studies focusing on L2 processing of different word orders and relying on the Interface Hypothesis were conducted on Spanish (Lozano, 2006, 2014; Lozano and Mendikoetxea, 2008, 2010; Dominguez and Arche, 2014). In online experiments, even advanced L2 learners were shown to have some vestigial difficulties with SV/VS orders. At the same time, corpus studies suggest that they understand the syntax-discourse aspects of VS structures, although they may have some problems with the grammatical representation of non-subject preverbal XPs in such sentences.

Another formal Second Language Acquisition (SLA) theory that may be relevant for our study is the Bottleneck Hypothesis (Slabakova, 2014). In the generative framework, syntax relies on universal principles, while morphology is highly idiosyncratic and language-specific. Accordingly, the Bottleneck Hypothesis predicts that mastering syntax is much easier than mastering morphology, which is the primary source of problems for L2 learners.

Successfully processing different word orders in Russian definitely depends on the knowledge of morphology, most notably, case morphology. And we know that L2 learners of Russian have problems with it both in production and in comprehension until the most advanced levels (Rubinstein, 1995a, 1995b; Cherepovskaia et al., 2021, 2022). However, we must admit that based on the very few existing studies, so far it is impossible to tell whether L2 problems with non-canonical word orders in Russian are syntactic or morphological in nature and to what extent.

A general problem that is discussed in many functional approaches to SLA is the role of L1: it was confirmed to affect even advanced L2 learners, especially in the domain of discourse (Rutherford, 1983; Green et al., 2000; Han, 2000; Jung, 2004). Many studies of word order focus on cross-linguistic differences in the domain of verb subcategorization: which arguments are encoded as subjects or objects, how often a particular verb is used as transitive or intransitive (e.g., Frenck-Mestre and Pynte, 1997; Witzel et al., 2012).

Russian can provide a lot of interesting material to test the hypotheses outlined above and to establish the relative importance of different factors. However, this was done in only one study so far, which is also the only study assessing context effects on the L2 processing of different word orders. Laleko (2022) analyzed the role of information structure and predicate-argument structure in the processing of canonical and non-canonical orders for three groups of participants: native speakers, heritage speakers (low and high proficiency) and adult learners of Russian.

The study involved assessing the acceptability of SV (O) and (O) VS sentences in different contexts, i.e., unlike most studies discussed above, it did not use online measures. Three types of predicates were used: transitive, unergative and unaccusative verbs [for unaccusative verbs, VS is the neutral word order, see also (Slioussar, 2011b)]. Contexts were such that target sentences either had a broad focus (all information was new), as in (4a), or a narrow subject focus, as in (4b). After each context sentence, two target sentences with different word orders were presented, as in (4c-d), and participants were asked to rate both of them on a 1 to 5 scale.

- (4) a. *Čto slučilos'?*
'What happened?'
b. *Kto počinil velosiped?*
'Who fixed the bicycle?'
c. *Papa počinil velosiped.* (SVO).
dad_{NOM.SG} fixed bicycle_{ACC.SG}.
d. *Velosiped počinil papa.* (OVS).
bicycle_{ACC.SG} fixed dad_{NOM.SG}.

Heritage and L2 speakers gave (O) VS structures lower ratings than native speakers. With SV and VS orders, information structure did not play a role for non-native speakers, but heritage speakers in the higher proficiency group were sensitive to the distinction between unaccusative and unergative verbs, like native speakers. With transitive verbs, higher proficiency heritage speakers demonstrated a native-like contrast in their ratings of OVS sentences with broad and narrow focus. Presumably, a given object may be a stronger trigger to use a non-canonical order than given information associated with the verb.

1.4 The present study

The goal of the present study was to compare L1 and L2 online and offline processing of different word orders in Russian. Our L2 participants were speakers of Mandarin Chinese. We chose SVO and OVS orders to have a canonical order and a non-canonical order with well-known information-structural properties and an inverted order of arguments, which is not characteristic for Chinese. Moreover, it was examined in several previous studies. In Experiment 1, target sentences were presented in isolation, while in Experiment 2, we used one-sentence contexts introducing one NP mentioned in the target sentence, like in several previous L1 studies (Kaiser and Trueswell, 2004; Slioussar, 2011a; Gattei et al., 2021).

Contexts introducing the first NP in the target sentence created a 'given – new' word order in it, which is characteristic for Russian, Chinese and many other languages with flexible word order (and, to a certain extent, to narrative texts universally). They can be viewed as appropriate. Contexts introducing the second NP created a 'new – given' word order in the target sentence and violated the information-structural requirements of OVS sentences (as we explained in section 1.1, SVO sentences are more flexible in this respect). They can be viewed as inappropriate. We aimed to find out how the word order factor and the context factor interact in L2 processing compared to L1 processing — a question that has been addressed in very few previous studies (and none of them compared appropriate and inappropriate contexts). This question was addressed in Experiment 2, while Experiment 1 examining the word order factor without the context factor can be seen as ancillary.

In both experiments, we measured word-by-word reading times to investigate online processing. After every sentence, we asked questions revealing whether readers interpreted it correctly, i.e., understood its predicate-argument structure. Finally, we also asked native speakers to evaluate how naturally target sentences sound on a 1 to 5 scale, tapping into their offline sensitivity to contextual requirements. Unlike Laleko (2022), we did not use this task with L2 participants (in her study, they were not sensitive to information-structural requirements of different word orders, only advanced heritage speakers were).

2 Experiment 1

In this experiment, our goal was to compare how native Russian speakers and Chinese learners process Russian sentences with different word orders (canonical SVO and inverted OVS) out of context and how native speakers evaluate them.

2.1 Participants

Two groups volunteered to take part in the study. The L1 group included 40 native Russian speakers (31 females) aged 18–43 (mean age 28.8). The L2 group consisted of 39 speakers of Mandarin Chinese (24 females) aged 18–35 (mean age 22.0). The experiment was carried out in accordance with the Declaration of Helsinki and existing Russian and international regulations concerning ethics in research. All participants provided informed consent. They received no financial reward for their participation.

All Chinese participants were students at Saint Petersburg State University in Russia and at the Belarusian State University in Belarus.⁴ They studied Russian at the Language testing center and at the Faculty of Philology of Saint Petersburg State University, and at the preparatory department of Belarusian State University. In total, 17 students were involved in different preparatory programs, 14 were in their first or second year of undergraduate studies, and 8 were in their third year. Twenty-three students had been living in Russia or Belarus for less than 1 year, 16 — for less than 2 years. When asked about their proficiency level in the Russian language, 19 people indicated the basic level (A2), 20 people — the lower intermediate, or the first certification level (B1). When asked about their proficiency in other foreign languages, all participants mentioned that they had studied English. Having more participants would be optimal, but the L2 groups we had access to (with a certain L1, a certain proficiency level etc.) were limited, unfortunately.

2.2 Materials

We constructed 16 sets of target sentences. Every set included two sentences that were identical except for the word order (SVO or OVS). Examples are given in (5a–b). We avoided object experiencer psych verbs or other constructions in which non-canonical orders may be more frequent than the canonical one (these verbs were discussed in the section 1.2). Since all sentences in the experiment were presented to participants segment-by-segment, we indicate how they were divided into segments.

- (5) a. *Russkij prepodavatel' / slušaet / kitajskogo studenta/i smotrit v okno.*
 Russian_{NOM.SG} teacher_{NOM.SG} / listens / Chinese_{ACC.SG} student_{ACC.SG} / and looks in window.

- b. *Kitajskogo studenta / slušaet / russkij prepodavatel' / i smotrit v okno.*
 Chinese_{ACC.SG} student_{ACC.SG} / listens / Russian_{NOM.SG} teacher_{NOM.SG} / and looks in window.
 'A / the Russian teacher is listening to a / the Chinese student and looking out the window.'

Thus, each sentence consisted of the following four segments:

- a subject NP (an animate noun in nominative singular with a preposed adjective);
- an object NP (an animate noun in accusative singular with a preposed adjective);
- a transitive verb in the present or past tense;
- the final segment (a second coordinated VP, a PP depending on the first verb etc.).

As we mentioned in the introduction, animacy may affect word order in the absence of information-structural differences. Therefore, subject and object NPs were balanced with respect to the animacy scale (both denoted either humans or animals). We chose only animate nouns to avoid forms with case syncretism (in most inanimate nouns, accusative forms coincide with nominative ones). The segments containing the object and the subject, which were crucial for our study, always consisted of two words. This was done to make reading time differences more pronounced. The final segment was introduced so that subject and object segments were not sentence-final.

We made sure that in all target sentences, it was impossible to guess grammatical roles of the NPs based on the semantics alone. For example, if (5a–b) are considered, both the teacher can look at the student and vice versa. Therefore, participants had to rely on case information to interpret these sentences correctly. To assess their interpretation accuracy, we constructed two questions for each target sentence set directed at the subject and at the object, like in (6a–b). All questions contained only a question word and a verb — we did not want to give our participants any further hints or to confuse them any further by adding any NPs.

- (6) a. *Kto slušaet?*
 who_{NOM.SG} listens.
 'Who is listening?'
 b. *Kogo slušajut?*
 whom_{ACC.SG} listen.
 'Who is being listened to?'

Rather than giving participants a choice of two answers, as it is usually done, we provided them with a window to type in their answer. This made the task more difficult for the L2 group, but we wanted to avoid guessing. The instructions before the experiment specified that brief answers (only the noun) were acceptable. Every participant saw an equal number of target

⁴ The teaching at the Belarusian State University is in Russian, and the absolute majority of people in Minsk where it is located also speak Russian as their native language. The Belarusian language is more widespread in other places in Belarus.

⁵ The verb is in the 3rd person plural form that can be used when the subject remains unspecified. It was important for us that in the question about the subject, the *wh*-word is in nominative, and in the question about the object, it is in accusative, and both questions do not contain any other NPs.

sentences in the two experimental conditions (SVO and OVS) with an equal number of subject and object questions.

Thus, every participant read 16 target sentences in one of the two conditions. We also constructed 10 filler sentences that were more syntactically diverse than target sentences to distract participants' attention from the experimental manipulation. Two examples are given in (7a-b). The questions for the fillers were directed at the PPs with a temporal or locative meaning. Filler sentences and questions were the same for every participant.

- (7) a. *K našemu deduške redko prihodjat raznye gosti.*
 to our_{DAT.SG} grandfather_{DAT.SG} rarely come various_{NOM.PL}
 guests_{NOM.PL}.
 b. *Bednyj xudožnik uexal iz Peterburga v pjatnicu večerom.*
 poor_{NOM.SG} artist_{NOM.SG} left from Petersburg_{GEN.SG} on
 Friday_{ACC.SG} at-night.

When constructing target and filler sentences, we selected vocabulary and grammatical features in accordance with the lexical minima and state standards for Russian as a foreign language. We made sure that they did not exceed the basic level according to the Russian State Testing System. Additionally, all sentences were checked on the online platform *Textometr*.⁶ They were generally assessed as being at the A1 level (elementary), and the A2 lexical list covered 87% of the vocabulary.

2.3 Procedure

The experiment was run on the web-based platform PCibex.farm. Data were collected in the presence of the experimenter or the Russian teacher of the L2 participants. We created two versions of the experiment for the L1 and L2 groups. In both groups, we measured sentence reading times and question answering accuracy. After that, the L1 group received a second task (evaluating target and filler sentences on a scale), while the L2 group was asked to fill in a questionnaire (about their native and foreign languages, and about their Russian studies in particular). For the L2 group, the experimental instructions and the questionnaire were translated into Chinese and checked by a native Chinese speaker.

To measure reading times, we used the moving window self-paced reading task (Just et al., 1982). Each trial began with a screen presenting a sentence, in which the words were masked by dashes, while spaces and punctuation remained intact. Each time the participant pressed the space bar, a segment was revealed, the previous segment was re-masked, and RTs were measured. After each sentence, a question and a window to type in the answer appeared. Participants were instructed to read at their natural pace. Two practice items were presented before the beginning of the experiment (in particular, we made sure that L2 readers understood the questions by giving them feedback).

In the second part of the experiment, L1 participants were asked to evaluate a number of sentences on the 1 to 5 scale, where 5 indicated a sentence that sounded fully natural in Russian, and 1

indicated a sentence that sounded completely unnatural. We included all target sentences from the first part in this task, as well as four filler sentences in which we modified the word order in a way that is not characteristic for Russian (although grammatical). We were interested to find out whether L1 participants subjectively perceive non-canonical word orders as sounding less natural than canonical ones in zero context. Sentences were shown on the screen one by one (unmasked). Before the main session, two practice items were presented.

2.4 Analysis

We analyzed participants' reading times, question answering accuracy and sentence ratings (in the L1 group). We did not analyze response times (these data were too noisy because our participants had to type their answers). Data from the two groups were analyzed separately.

During the preliminary data processing in the L1 group, RTs that exceeded a threshold of 2.5 standard deviations, by segment and by condition, were excluded as outliers (Ratcliff, 1993). In total, this led to the exclusion of 4.8% of the data. There was no filtering based on accuracy since all participants performed well, providing over 86% correct responses.

In the L2 group, the task appeared to be too difficult for many participants, as we could judge from their low accuracy. To have an exclusion criterion independent from our experimental manipulations, we discarded data from participants who scored below 60% correct on questions to filler sentences. As a result, data from 26 out of 39 participants were included in the final analysis. Subsequently, 5.0% of the RTs were excluded because they exceeded a threshold of 2.5 standard deviations, by segment and by condition.

The statistical analysis was done in the R programming environment.⁷ We modeled RT data with a mixed-effects regression using the *lmer* function from the *lme4* package, accuracy data with a mixed-effects logistic regression using the *glmer* function from the *lme4* package, and sentence rating data with a mixed-effects ordinal regression using the *glmer* function from the *lme4* package (Bates et al., 2015). To obtain the *p* values from the *t* values given by the model, we used the *lmerTest* package (Kuznetsova et al., 2015). Random intercepts and random slopes by a participant and by an item were included in the models.

In the analysis of sentence ratings, word order (SVO or OVS) was the only fixed effect. In the analysis of RTs, we added segment length. Some NPs and verbs that we used were longer than the others, so it was significant in most comparisons, as expected. But this variation could have been covered by random effects, so this factor was not interesting to us *per se*. The reason to include it was that some accusative singular forms of nouns and adjectives are one letter longer than nominative forms, and we wanted to make sure that if there are any differences between the two experimental conditions, they cannot be reduced to that. In the analysis of answering accuracy, we used the word order factor and two factors

⁶ <https://textometr.ru/>

⁷ www.r-project.org

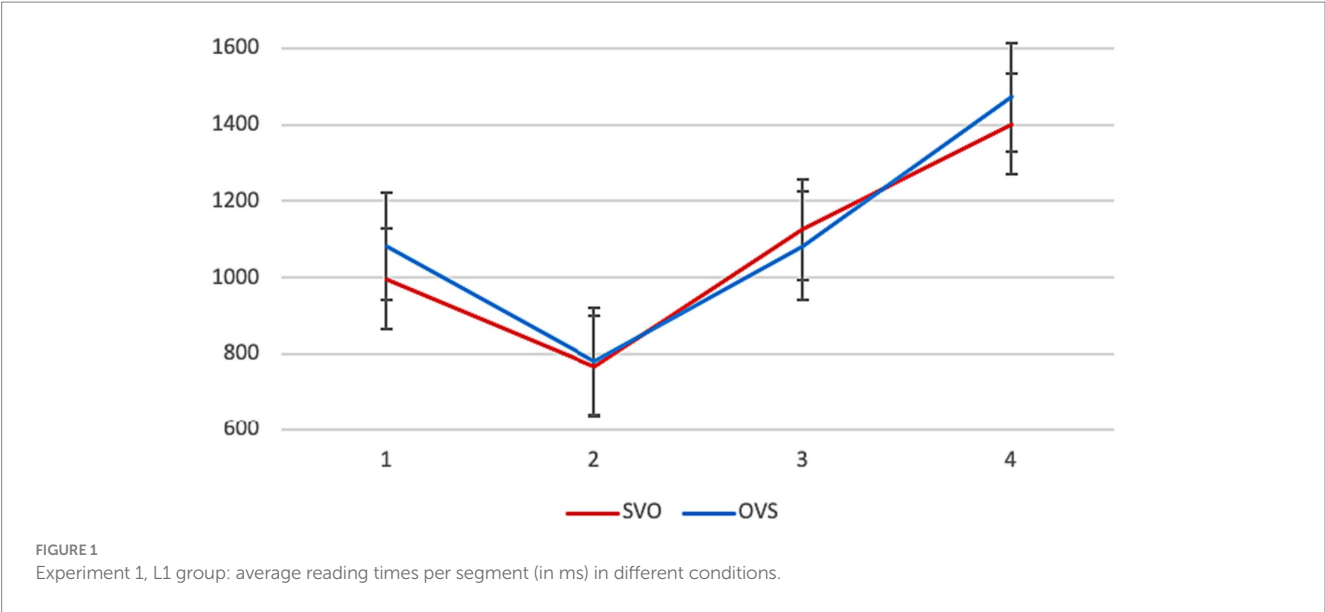


TABLE 1 Experiment 1, L1 group: the proportion of correct answers in different conditions.

Word order	NP role	NP position	Correct answers
SVO	S	1	93%
SVO	O	3	92%
OVS	O	1	87%
OVS	S	3	86%

capturing to which NP the question was directed: to the subject or the object (NP role) and to the first or the third segment (NP position). A preliminary examination of the data from the L2 group suggested that the later factor may be important, and the subsequent statistical analysis confirmed that.

2.5 Results

2.5.1 L1 group: reading times

Average reading times in different conditions are presented in Figure 1. No significant differences between the two word orders were found.⁸

2.5.2 L1 group: question answering accuracy

The average proportion of correct answers by condition is presented in Table 1. Out of three factors we analyzed, the word order proved to be significant. Overall, L1 participants performed very well, but made more errors with OVS sentences ($\beta = -0.78$, $SE = 0.31$, $p = 0.013$).

TABLE 2 Experiment 1, L1 group: average ratings of sentences in different conditions.

Word order	Rating
SVO	4.48
OVS	3.05

2.5.3 L1 group: sentence ratings

Average ratings of SVO and OVS sentences are presented in Table 2. The SVO order was rated significantly higher than the OVS one ($\beta = -2.65$, $SE = 0.75$, $p < 0.001$).

2.5.4 L2 group: reading times

Average reading times are presented in Figure 2. Significant differences between the two conditions were found in the first segment. The first NP is read faster when it is the subject (in SVO) than when it is the object (in OVS) ($\beta = 726.80$, $SE = 290.77$, $p = 0.013$).⁹

2.5.5 L2 group: question answering accuracy

The average proportion of correct responses by condition is given in Table 3. Two factors were significant: the word order and the NP position, while the NP role was not. It was easier for L2 participants to answer questions about SVO sentences ($\beta = -0.56$, $SE = 0.21$, $p = 0.011$) and about the first NP in the sentence ($\beta = -0.92$, $SE = 0.22$, $p < 0.001$). Maybe, this NP was better memorized. Another possibility was suggested by an anonymous reviewer. In Chinese, *wh*-phrases do not have a designated position in the beginning of the sentence, i.e., they stay *in situ*, so subject *wh*-phrases are preverbal, like NP subjects, while object *wh*-phrases follow the verb, like NP objects. Maybe, our participants sometimes treated all *wh*-words and all preverbal NPs as

⁸ The length factor was significant in the first, second and fourth segments ($\beta = 31.72$, $SE = 9.89$, $p = 0.005$; $\beta = 23.89$, $SE = 7.96$, $p = 0.009$; $\beta = 47.41$, $SE = 9.67$, $p < 0.001$).

⁹ The length factor was significant in the second segment ($\beta = 214.60$, $SE = 38.54$, $p < 0.001$).

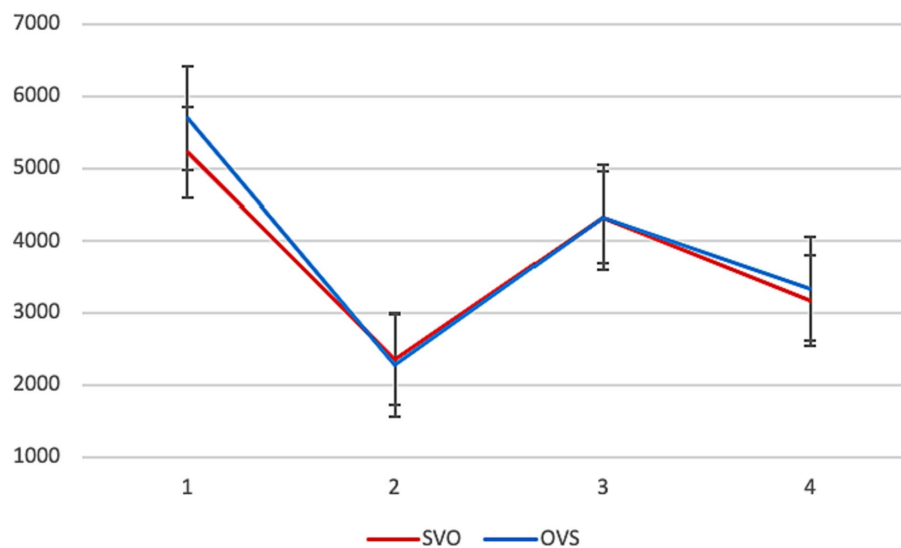


FIGURE 2
Experiment 1, L2 group: average reading times per segment (in ms) in different conditions.

TABLE 3 Experiment 1, L2 group: the proportion of correct answers in different conditions.

Word order	NP role	NP position	Correct answers
SVO	S	1	74%
SVO	O	3	46%
OVS	O	1	55%
OVS	S	3	43%

subjects, which gave them a chance to respond correctly when neither of these assumptions was correct.

2.6 Discussion

The goal of this experiment was to examine how isolated sentences with different word orders are processed by native speakers of Russian and by learners of Russian as a foreign language. In online processing, the OVS order did not pose any significant difficulties for L1 readers, but interpretation accuracy for OVS sentences was slightly lower. This is expected given that sentences with non-canonical orders are widespread in Russian, although the canonical order is still by far the most frequent. However, word order alternations are regulated by information structure, so non-canonical orders have certain contextual requirements and are not used in isolation. We demonstrated that native speakers are sensitive to that when they evaluate how natural sentences with different orders sound to them.

The picture is different for L2 readers. In OVS sentences, they slow down on the first segment, when it becomes clear that they are dealing with a non-canonical word order. Their interpretation accuracy shows that online difficulties often result in the ultimate failure to construct a correct interpretation, in particular, to understand the predicate-argument structure of the sentence. The fact that L2 readers answer questions about the first NP in

the sentence more accurately also stresses that processing several arguments and understanding their semantic roles is difficult for them. Difficulties associated with understanding *wh*-questions may aggravate the situation. To tease apart these factors, one may turn to a different experimental design, in which participants are asked to choose a picture that corresponds to a sentence rather than to answer *wh*-questions.

At the same time, let us note that we do not see any significant differences associated with case *per se* (it could be the case that accusative NPs had longer RTs in any position or triggered more interpretation errors). L2 readers resemble L1 readers in this respect, but the similarity may be deceiving. For L1 readers, processing case information is too easy to produce any noticeable effects. For our L2 group, it may be too difficult: maybe, we do not see any effects because they usually fail to do so, which is eventually reflected in their low question-answering accuracy. Further studies with more advanced L2 participants are necessary to find a definitive answer.

Finally, let us note the following difference between L1 and L2 groups. L2 participants take the longest to read the first segment, while RTs for the final segment, which reflect late stages of syntactic processing, are relatively short compared both to the first and the third segment. In contrast, L1 readers process NPs in the first and the third segments relatively fast and slow down on the last segment — presumably, to complete the syntactic representation of the sentence. Judging by their low accuracy, L2 participants often skip this step, being overloaded with syntactic processing, and, consequently, fail to arrive at the correct interpretation of the sentence.

3 Experiment 2

In this experiment, we aimed to compare how native Russian speakers and Chinese learners of Russian process sentences with different word orders (SVO and OVS) in the contexts satisfying or not satisfying their information-structural requirements. We also tested how native speakers evaluate them.

3.1 Participants

Like in Experiment 1, there were two groups of participants. The L1 group included 51 native Russian speakers (38 females) aged 17–47 (mean age 27.0). The L2 group consisted of 44 speakers of Mandarin Chinese (27 females) aged 18–25 (mean age 21.4). The experiment was carried out in accordance with the Declaration of Helsinki and existing Russian and international regulations concerning ethics in research. All participants provided informed consent and volunteered to participate without any financial reward.

All Chinese participants studied Russian at the Language testing center or at the Faculty of Philology of Saint Petersburg State University. In total, 5 students were involved in preparatory programs, 14 were in their first or second year of undergraduate studies, and 25 were in their third year. Sixteen students had been living in Russia for less than 1 year, 15 — for less than 2 years, 13 — for less than 3 years. When asked about their proficiency level in the Russian language, 9 students indicated the basic level (A2), and 35 people — the lower intermediate, or the first certification level (B1). Additionally, they all mentioned that they also studied English as a foreign language.

3.2 Materials

We took 16 target sentence sets from Experiment 1 (including questions) and constructed two one-sentence contexts for them. Context sentences always mentioned the subject or object from the target sentence. Examples are given in (8a–b) (in (9a–b), we repeat examples of target sentences given in (5a–b) above).

- (8) a. *Russkij prepodavatel' / provodit / zanjatie v auditorii.*
 Russian_{NOM.SG} teacher_{NOM.SG} / conducts / lesson in classroom.
 'A Russian teacher is conducting a lesson in the classroom.'
 b. *Kitajskij student / prišel / na zanjatie v auditoriyu.*
 Chinese_{NOM.SG} student_{NOM.SG} / came / to lesson in classroom.
 'A Chinese student came to a lesson in the classroom.'
- (9) a. *Russkij prepodavatel' / slušaet / kitajskogo studenta / i smotrit v okno.*
 Russian_{NOM.SG} teacher_{NOM.SG} / listens / Chinese_{ACC.SG} student_{ACC.SG} / and looks in window.
 b. *Kitajskogo studenta / slušaet / russkij prepodavatel' / i smotrit v okno.*
 Chinese_{ACC.SG} student_{ACC.SG} / listens / Russian_{NOM.SG} teacher_{NOM.SG} / and looks in window.
 'The Russian teacher is listening to the Chinese student and looking out the window.'

If we present (9a) after (8a) and (9b) after (8b), target sentences will start with a given NP followed by a new one (we will term this *G-N contexts*). This is characteristic for languages with a free word order, including Russian, so we can consider G-N contexts appropriate for the respective target sentences, or satisfying their information-structural requirements. If we present (9a) after (8b) and (9b) after (8a), target sentences will start with a new NP followed by a given one (*N-G contexts*). N-G contexts are infrequent in Russian and can be found only in special constructions like focus fronting. In our case, no focus fronting can

be expected, so these contexts can be considered inappropriate, or not satisfying the information-structural requirements of target sentences. As we noted in the introduction, the canonical SVO order is compatible with a wider range of contexts, while other orders, like OVS, have much stricter context requirements. In this study, we aim to find out whether participants are sensitive to N-G contexts in general and to the contextual requirements of different orders.

Context sentences satisfied the same requirements for vocabulary and grammar as target sentences did. The character from the target sentence was mentioned at the beginning or in the middle of the context sentence to give readers some time to accommodate this information. We used the same NP as in the target sentence (to leave no room for confusion) in the nominative singular form. This is a potential limitation of our study that can be addressed in further research: for L2 participants, it may be easier to read target sentences in which the given NP is in the same case as in the context sentence. Pairs of context sentences could be different in the beginning, but the end was always the same to avoid any effects in the following target sentence. We made sure that context sentences do not provide any hints on the distribution of grammatical roles in target sentences.

We also took 10 filler sentences with questions from Experiment 1 and created context sentences for them. These context sentences did not vary and could be considered appropriate (G-N). In the second part of the experiment, in which L1 participants rated sentences, four filler sentences with a modified word order were used, like in Experiment 1. Since the word order changed, the context became N-G (inappropriate).

3.3 Procedure

The procedure was the same as in the Experiment 1.

3.4 Analysis

Like in Experiment 1, we analyzed participants' reading times in target sentences (context sentence data were not included in the analysis), question answering accuracy and sentence ratings (in the L1 group). Data from the two groups were analyzed separately. During the preliminary data processing, we excluded 12 out of 44 L2 participants who gave less than 60% correct answers to the questions to filler sentences. In the L1 group, all participants answered more than 85% questions correctly. Then RTs that exceeded a threshold of 2.5 standard deviations, by segment and by condition, were removed (3.7% of the data in the L1 group, 5.4% in the L2 group).

The statistical analysis was the same as in Experiment 1. In the analysis of RTs and ratings, two fixed effects were included: word order (SVO or OVS) and context (G-N or N-G). In the discussion section, we will come back to the question which context effects can be explained by its global and local properties (i.e., by its general (in)appropriateness or by givenness/newness of certain NPs).

As for answering accuracy, we should note that due to a technical issue, there was a problem with the design: questions related to the subject were always asked after sentences in a G-N context, while questions related to the object were asked after sentences in a N-G context. Thus, the NP role factor (subject or object) was coupled

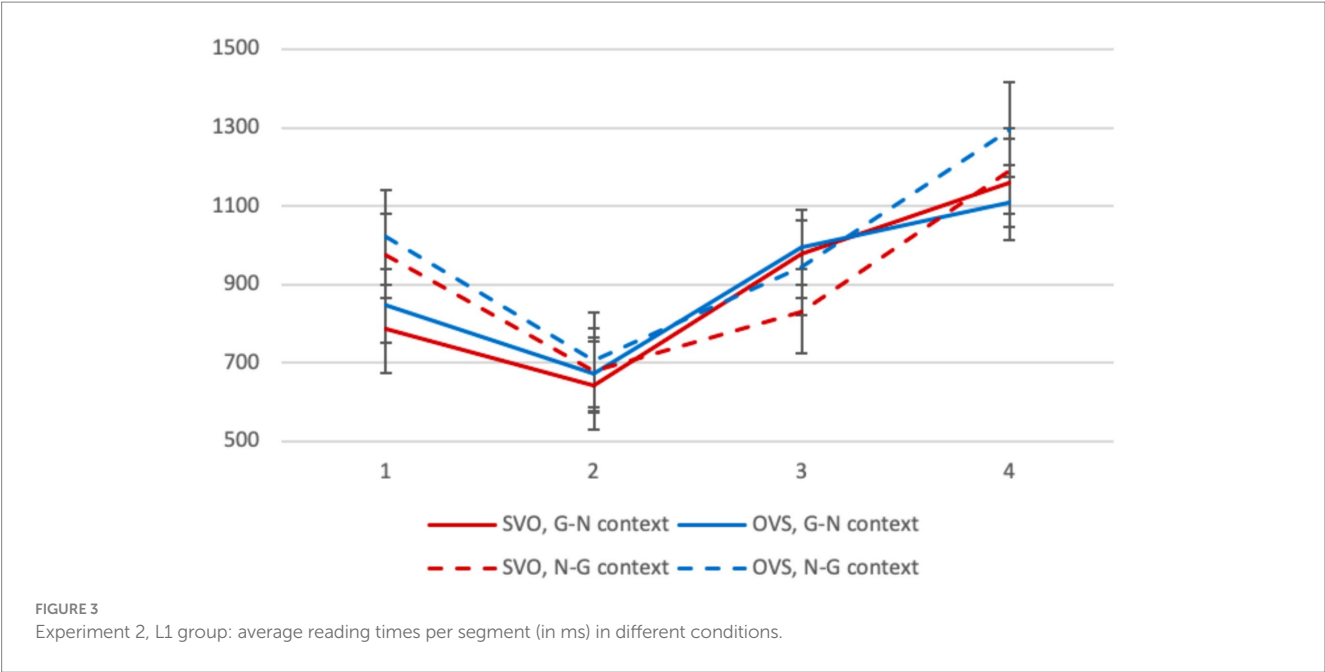


TABLE 4 Experiment 2, L1 group: the proportion of correct answers in different conditions.

Context	Word order	NP role	NP position	NP givenness	Correct answers
G-N	SVO	S	1	given	91%
N-G	SVO	O	3	given	88%
G-N	OVS	S	3	new	88%
N-G	OVS	O	1	new	86%

with the context factor, and the word order factor was coupled with NP givenness. The NP position factor remained independent. Thus, we can interpret the obtained results only with significant limitations, but will nevertheless propose an interpretation in the discussion section.

3.5 Results

3.5.1 L1 group: reading times

Average reading times in different conditions are presented in Figure 3. On the first segment, the context factor was significant: sentences in the G-N context were read faster ($\beta = 181.79$, $SE = 30.61$, $p < 0.001$). Of course, this effect may be explained, at least partially, by the fact that given NPs are read faster than new ones (especially given the fact that they were literally repeated). There were no significant differences between conditions on the second segment.¹⁰

On the third segment, the context factor was significant, and this was definitely due to NP givenness: given NPs in N-G contexts were read faster ($\beta = -145.70$, $SE = 34.53$, $p < 0.001$). The interaction between the two factors reached significance as well: OVS sentences

TABLE 5 Experiment 2, L1 group: average ratings of sentences in different conditions.

Context	Word order	Rating
G-N	SVO	4.44
N-G	SVO	3.84
G-N	OVS	4.41
N-G	OVS	3.06

in the N-G context were read slower than SVO ones ($\beta = 106.49$, $SE = 48.94$, $p = 0.030$). This cannot be explained by the local properties of NPs and is most probably due to the fact that OVS sentences have much stricter context requirements, and native speakers are sensitive to that. On the final segment, there was a significant interaction between the two factors, similar to that observed for the third segment ($\beta = 168.85$, $SE = 83.20$, $p = 0.043$). The effect of context was visible for OVS sentences, but not for SVO ones.

3.5.2 L1 group: question answering accuracy

The average proportion of correct answers by condition is presented in Table 4. L1 participants performed very well, and no differences reached significance, although certain tendencies can be seen.

3.5.3 L1 group: sentence ratings

Average ratings of sentences in different conditions are presented in Table 5. G-N contexts were rated significantly higher ($\beta = -1.62$,

¹⁰ The length factor was significant in the second, third and fourth segments ($\beta = 9.46$, $SE = 4.30$, $p = 0.045$; $\beta = 28.51$, $SE = 5.76$, $p < 0.001$; $\beta = 34.15$, $SE = 4.05$, $p < 0.001$).

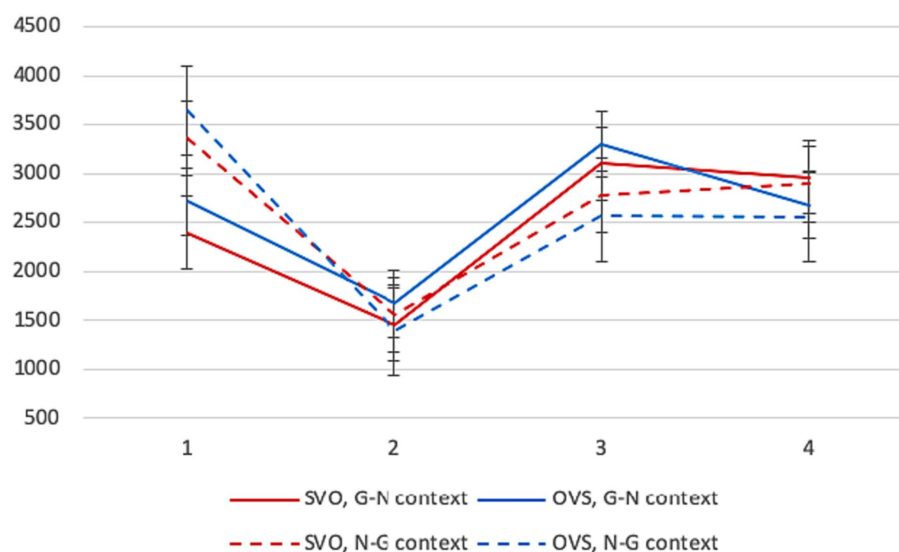


FIGURE 4

Experiment 2, L2 group: average reading times per segment (in ms) in different conditions.

$SE = 0.21, p < 0.001$). The interaction was also significant: like with the RT data, the effect of context was more pronounced for OVS sentences ($\beta = -1.54, SE = 0.29, p < 0.001$).

3.5.4 L2 group: reading times

Average reading times in different conditions are presented in Figure 4. The context factor was highly significant on the first segment, like for L1 participants. Given NPs in the G-N context were read much faster ($\beta = 945.92, SE = 225.42, p < 0.001$). The word order factor reached significance as well: SVO was easier ($\beta = 389.94, SE = 230.35, p = 0.032$).¹¹

There were no significant differences between conditions on the second segment. On the third segment, the context factor was significant: given NPs in N-G contexts were read faster ($\beta = -572.29, SE = 208.41, p = 0.006$). This result was similar to the L1 group, but other patterns were not. The interaction between the context and word order factors reached significance: the effect of givenness was more pronounced for OVS sentences ($\beta = -359.15, SE = 292.75, p = 0.041$). Notably, L1 participants read OVS sentences in N-G contexts more slowly than SVO sentences, being sensitive to their stricter context requirements. We do not see this sensitivity in L2 readers who show the opposite pattern. No differences between conditions reached significance on the final segment.

3.5.5 L2 group: question answering accuracy

The average proportion of correct answers by condition is presented in Table 6 and summarized in Table 7. As we noted above, due to a technical issue, questions related to the subject were always asked after G-N sentences, while questions related to the object — after N-G sentences. Therefore, some factors were coupled, and we should try to tease them apart when interpreting the results.

The context type and NP role factors, which did not reach significance, could not cancel each other out because they were supposed to work in the same direction. The NP role did not affect answering accuracy in Experiment 1 or RTs in both experiments. The effects of context on RTs in the L2 group are mostly local. L2 participants find given NPs easier to read (which was captured by the NP givenness factor in the current analysis), but are not sensitive to the global (in)appropriateness of the context.

3.6 Discussion

The goal of the second experiment was to compare how sentences with different word orders are processed by L1 and L2 participants in different contexts: G-N and N-G. Similarly to Experiment 1, the word order factor significantly affected RTs and question-answering accuracy only in the L2 group. For L1 participants, processing non-canonical orders was not particularly difficult.

The context factor played a major role both for L1 and for L2 participants. As we noted above, its effects can be explained locally (by the givenness of particular NPs) or globally (by the fact that G-N contexts are characteristic for Russian and can be seen as appropriate, while N-G contexts are not). A slowdown associated with it was much larger on the first segment (on new NPs in the N-G context) than on the third segment (on new NPs in the G-N context) in both groups. We can conclude that for all readers, both local and global aspects are important, although local ones play a larger role.¹² Sentence ratings in the L1 group can be affected only by the

¹¹ The length factor was significant in the second and fourth segments ($\beta = 133.91, SE = 24.44, p < 0.001$; $\beta = 70.90, SE = 22.04, p = 0.006$).

¹² Most probably, they also affected accuracy in the L2 group, but their effects cannot be teased apart from the word order effects (in the L1 group, no factor reached significance in this task).

TABLE 6 Experiment 2, L2 group: the proportion of correct answers in different conditions.

Context	Word order	NP role	NP position	NP givenness	Correct answers
G-N	SVO	S	1	given	80%
N-G	SVO	O	3	given	62%
G-N	OVS	S	3	new	37%
N-G	OVS	O	1	new	46%

TABLE 7 Experiment 2, L2 group: the proportion of correct answers depending on different factors.

Word order + NP givenness		Context / NP role		NP position	
SVO + given	OVS + new	G-N + S	N-G + O	1	3
71%	42%	59%	54%	63%	50%

Two groups of factors significantly affected the results: the word order/givenness ($\beta = -1.37$, $SE = 0.20$, $p < 0.001$) and the NP position ($\beta = -0.70$, $SE = 0.20$, $p < 0.001$). In case of word order and givenness, both factors probably influenced the results working in the same direction and resulting in the most pronounced effect (the former was significant in Experiment 1 and for RT data in the current experiment, the latter was the main factor affecting RTs). The NP position factor that played a role in Experiment 1 remained significant, as expected.

global (in)appropriateness of the context, and we can see that this factor was significant.

However, while the global RT picture is similar for L1 and L2 participants, there are also some principled differences. For L1 readers, the effect of the inappropriate N-G context is more pronounced for OVS orders. This is evident not only in RTs (given subjects following new objects are read almost as slowly as new NPs), but also in sentence ratings. This can be explained by more strict context requirements for non-canonical orders. L2 readers do not exhibit a similar subtle sensitivity to context.

Finally, let us note that NP position affected accuracy in the L2 group, like in Experiment 1. It was easier to answer questions about the first NP in the sentence. Given that the overall accuracy was low, this supports the conclusion that L2 readers have problems with processing several arguments and understanding their semantic roles. Problems with understanding *wh*-questions could also contribute to this.

4 Conclusion

In this study, we compared how native speakers of Russian and speakers of Mandarin Chinese learning Russian as a foreign language process Russian sentences with different word orders in isolation (Experiment 1) and in context (Experiment 2). We chose SVO and OVS sentences for the comparison to have a canonical order and a non-canonical order with well-known information-structural properties and an inverted order of arguments, which is not characteristic for Chinese. One-sentence contexts introduced one NP mentioned in the target sentence, either the first or the second. Thus, in the former case, given information preceded new information in the target sentence, which is characteristic for Russian and many other languages, while in the latter case, the opposite was true. We used different measures to capture online and offline effects: word-by-word reading times, question-answering accuracy and sentence rating on a 1 to 5 scale (for L1 participants).

In both experiments, RTs and question-answering accuracy data showed that non-canonical orders were difficult for L2 participants, but not for L1 participants (for them, the effects of this factor were small or absent altogether). However, L1 participants gave non-canonical orders

lower ratings in isolation, presumably because in naturally occurring texts, they are used only in particular contexts. It would be interesting to find out to what extent these difficulties are universal for L2 processing, or native speakers of other languages in which subject-object inversion is possible, like in Russian, would not experience them. Further research may also focus on other non-canonical word orders. For example, would SOV be more difficult than SVO for L2 readers, or only changing the relative order of arguments creates substantial problems? How would SV(XP) vs. (XP)VS orders with intransitive verbs, like the ones examined by Laleko (2022), be processed?

As for the context factor in Experiment 2, some effects are universal for L1 and L2 processing: all participants read given NPs faster than new ones and preferred sentences with a ‘given – new’ word order. The latter may reflect the universal principles of narrative coherence — then L2 readers do not need to acquire this knowledge, they only need to apply it to a new language. However, unlike native speakers, they are not sensitive to more subtle contextual requirements of different word orders, in particular, to the fact that the canonical word order is acceptable in a much wider range of contexts, while non-canonical orders heavily depend on the appropriate context to sound natural. These results are interesting to compare with those by Laleko (2022) who found that L2 learners are not sensitive to information-structural requirements when asked to evaluate SV(O) and (O)VS sentences. Thus, although it might be easier for them to process ‘given – new’ orders, this does not necessarily crystallize into knowledge how different orders should be used.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://osf.io/9eust/?view_only=d5205d3a3b7644e595e4910470e4f095.

Ethics statement

Ethical approval was not required for the studies involving humans because The local Russian legislation does not require ethics committee approval for behavioral studies with adult non-vulnerable

participants IF they provide written informed consent and IF the study is carried out in accordance with the Declaration of Helsinki and existing Russian and international regulations concerning ethics in research. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

NS: Conceptualization, Investigation, Project administration, Writing – original draft. MH: Data curation, Investigation, Writing – original draft.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. The study was

supported by the grant 75288744, 121050600033–7 from Saint Petersburg State University.

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Bader, M., and Meng, M. (1999). Subject-object ambiguities in German embedded clauses: an across-the-board comparison. *J. Psychol. Res.* 28, 121–143. doi: 10.1023/A:1023206208142
- Bailyn, J. F. (1995). A configurational approach to Russian “free” word order. PhD Dissertation, Cornell University.
- Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). *lme4: linear mixed-effects models using Eigen and S4*. R package version 1.1–8. Available at: <http://CRAN.R-project.org/package=lme4/> (Accessed September 15, 2023)
- Bornkessel, I., Schlesewsky, M., and Friederici, A. D. (2003). Contextual information modulates initial processes of syntactic integration: the role of inter-versus intrasentential predictions. *J. Exp. Psychol. Learn. Mem. Cogn.* 29, 871–882. doi: 10.1037/0278-7393.29.5.871
- Burmester, J., Spalek, K., and Wartenburger, I. (2014). Context updating during sentence comprehension: the effect of aboutness topic. *Brain Lang.* 137, 62–76. doi: 10.1016/j.bandl.2014.08.001
- Callies, M. (2009). *Information highlighting in advanced learner English. The syntax-pragmatics interface in second language acquisition*. Amsterdam: John Benjamins
- Cherepovskaia, N., Reutova, E., and Slioussar, N. (2021). Becoming native-like for good or ill: online and offline processing of case forms in L2 Russian. *Front. Psychol.* 12:652463. doi: 10.3389/fpsyg.2021.652463
- Cherepovskaia, N., Slioussar, N., and Denissenko, A. (2022). Acquisition of the nominal case system in Russian as a second language. *Second. Lang. Res.* 38, 555–580. doi: 10.1177/0267658320988058
- Clahsen, H., and Featherston, S. (1999). Antecedent priming at trace positions: evidence from German scrambling. *J. Psychol. Res.* 28, 415–437. doi: 10.1023/A:1023293132656
- DeKeyser, R. M. (2005). What makes learning second language grammar difficult? A review of issues. *Lang. Learn.* 55, 1–25. doi: 10.1111/J.0023-8333.2005.00294.X
- Dominguez, L., and Arche, M. J. (2014). Subject inversion in non-native Spanish. *Lingua* 145, 243–265. doi: 10.1016/J.LINGUA.2014.04.004
- Dryer, M. (2005). Order of subject, object, and verb. The World Atlas of Language Structures. Available at: <https://wals.info/chapter/81> (Accessed October 15, 2023).
- Erdocia, K., Laka, I., Mestres-Misse, A., and Rodriguez-Fornells, A. (2009). Syntactic complexity and ambiguity resolution in a free word order language: behavioral and electrophysiological evidences from Basque. *Brain Lang.* 109, 1–17. doi: 10.1016/j.bandl.2008.12.003
- Frazier, L., and Flores d'Arcais, G. B. (1989). Filler-driven parsing: a study of gap-filling in Dutch. *J. Mem. Lang.* 28, 331–344. doi: 10.1016/0749-596X(89)90037-5
- Frenc-Mestre, C., and Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *Quarter. J. Exp. Psychol. Hum. Exp. Psychol.* 50, 119–148. doi: 10.1080/027249897392251
- Gao, Q. (2008). “Word order in mandarin: Reading and speaking” in *Proceedings of the 20th north American conference on Chinese linguistics* 2. eds. M. Chan and H. Kang (Columbus, OH: Ohio State University), 611–626.
- Gattei, C. A., Dickey, M. W., Wainelboim, A. J., and Paris, L. (2015). The thematic hierarchy in sentence comprehension: a study on the interaction between verb class and word order in Spanish. *Q. J. Exp. Psychol.* 68, 1981–2007. doi: 10.1080/17470218.2014.1000345
- Gattei, C. A., Paris, L. A., and Shalom, D. E. (2021). Information structure and word order canonicity in the comprehension of Spanish texts: an eye-tracking study. *Front. Psychol.* 12:629724. doi: 10.3389/fpsyg.2021.629724
- Gattei, C. A., Sevilla, Y., Tabullo, A., Wainelboim, A., Paris, L., and Shalom, D. (2017). Prominence in Spanish sentence comprehension: an eye-tracking study. *Lang. Cogn. Neurosci.* 33, 587–607. doi: 10.1080/23273798.2017
- Gibson, E. (1998). Linguistic complexity: locality of syntactic dependencies. *Cognition* 68, 1–76. doi: 10.1016/S0010-0277(98)00034-1
- Green, C. F., Christopher, E. R., Lam, J., and Mei, K. (2000). The incidence and effects on coherence of marked themes in interlanguage texts: a corpus-based enquiry. *Engl. Specif. Purp.* 19, 99–113. doi: 10.1016/S0889-4906(98)00014-3
- Han, Z. (2000). Persistence of the implicit influence of NL: the case of the pseudopassive. *Appl. Linguis.* 21, 78–105. doi: 10.1093/APPLIN/21.1.78
- Hyönä, J., and Hujanen, H. (1997). Effects of case marking and word order on sentence parsing in Finnish: an eye fixation analysis. *Q. J. Exp. Psychol.* 50, 841–858. doi: 10.1080/713755738
- Jung, E. H. (2004). Topic and subject prominence in interlanguage development. *Lang. Learn.* 54, 713–738. doi: 10.1111/J.1467-9922.2004.00284.X
- Just, M. A., Carpenter, P. A., and Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *J. Exp. Psychol. Gen.* 111, 228–238. doi: 10.1037//0096-3445.111.2.228
- Kaiser, E., and Trueswell, J. C. (2004). The role of discourse context in the processing of a flexible word-order language. *Cognition* 94, 113–147. doi: 10.1016/j.cognition.2004.01.002
- Kodzasov, S. V. (1996). “Kombinatornaja model' frazovoj prosodii (in Russian, 'A combinatorial model of phrasal prosody')” in *Prosodičeskij stroj russkoj reči (in Russian, 'prosodic system of the Russian language')* Ed. T. Nikolaeva (Moscow: Institute of the Russian Language RAS), 85–123.
- Kovtunova, I. I. (1976). *Sovremennij russkij jazyk. Porjadok slov i aktual'noe členenie predloženija (in Russian, 'Modern Russian language. Word order and information structure')*. Moscow: Prosveščenie.
- Krylova, O. A. (1992). *Kommunikativnyj sintaksis russkogo yazyka (in Russian, 'communicative syntax of the Russian language')*. Moscow: Russian University of Friendship of Peoples Publishing House.
- Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2015). *lmerTest: tests in linear mixed effects models*. R package version 2.0–25. Available at: <http://CRAN.R-project.org/package=lmerTest> (Accessed September 15, 2023).
- Laleko, O. (2022). Word order and information structure in heritage and L2 Russian: focus and unaccusativity effects in subject inversion. *Int. J. Bilingualism* 26, 749–766. doi: 10.1177/13670069211063674

- Li, C., and Thompson, S. (1974). An explanation of word order change: SVO > SOV. *Found. Lang.* 12, 201–214.
- Lozano, C. (2006). Focus and split intransitivity: the acquisition of word order alternations in non-native Spanish. *Second. Lang. Res.* 22, 145–187. doi: 10.1191/0267658306sr264oa
- Lozano, C. (2014). “Word order in second language Spanish,” in *The Handbook of Spanish Second Language Acquisition*. Ed. K. L. Geeslin. Hoboken, NJ: John Wiley & Sons. 287–310.
- Lozano, C., and Mendikoetxea, A. (2008). “Postverbal subjects at the interfaces in Spanish and Italian learners of L2 English: A corpus analysis,” in *Linking up contrastive and learner corpus research*. Eds. G. Gilquin, S. Papp and M. B. Diez-Bedmar (Leiden: Brill), 83–125.
- Lozano, C., and Mendikoetxea, A. (2010). Interface conditions on postverbal subjects: a corpus study of L2 English. *Biling. Lang. Cogn.* 13, 475–497. doi: 10.1017/S1366728909990538
- Miyamoto, E., and Takahashi, S. (2000). Effects of word-order variation on on-line sentence processing in Japanese. *J. Psychol. Res.* 29, 511–529.
- Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychol. Bull.* 114, 510–532. doi: 10.1037/0033-2909.114.3.510
- Rubinstein, G. (1995a). On acquisition of Russian cases by American classroom learners. *IRAL Int. Rev. Appl. Linguist Lang. Teach.* 33, 10–34. doi: 10.1515/iral.1995.33.1.9
- Rubinstein, G. (1995b). On case errors made in oral speech by American learners of Russian. *Slav. East Eur. J.* 39, 408–429. doi: 10.2307/308241
- Rutherford, W. (1983). “Language typology and language transfer” in *Language transfer in language learning*. eds. S. M. Gass and L. Selinker (Rowley, MA: Newbury House), 358–370.
- Schumacher, P. B., and Hung, Y.-C. (2012). Positional influences on information packaging: insights from topological fields in German. *J. Mem. Lang.* 67, 295–310. doi: 10.1016/j.jml.2012.05.006
- Sekerina, I. A. (1999). The scrambling complexity hypothesis and processing of split scrambling constructions in Russian. *J. Slav. Linguist.* 7, 265–304.
- Sekerina, I. A. (2003). “Scrambling and processing: dependencies, complexity and constraints” in *Word order and scrambling*. ed. S. Karimi (Oxford: Blackwell), 301–324.
- Sirotnina, O. B. (1965). *Porjadok slov v russkom jazyke (in Russian, ‘word order in the Russian language’)*. Saratov: Saratov State University Press.
- Slabakova, R. (2014). The bottleneck of second language acquisition. *Foreign Lang. Teach. Res.* 46, 543–559.
- Slioussar, N. (2007). Grammar and information structure. A study with reference to Russian. Doctoral dissertation, University of Utrecht.
- Slioussar, N. (2011a). Processing of a free order language: the role of syntax and context. *J. Psycholinguist. Res.* 40, 291–306. doi: 10.1007/s10936-011-9171-5
- Slioussar, N. (2011b). Russian and the EPP requirement in the tense domain. *Lingua* 121, 2048–2068. doi: 10.1016/j.lingua.2011.07.009
- Slioussar, N., and Makarchuk, I. (2022). SOV in Russian: a corpus study. *J. Slavic Linguist.* 30, 1–14.
- Sorace, A. (2011). Pinning down the concept of “interface” in bilingualism. *Linguist. Approach. Biling.* 1, 1–33. doi: 10.1075/LAB.1.1.01SOR
- Sorace, A. (2012). Pinning down the concept of “interface” in bilingualism: a reply to peer commentaries. *Linguist. Approach. Biling.* 2, 209–217. doi: 10.1075/LAB.2.2.04SOR
- Sorace, A., and Serratrice, L. (2009). Internal and external interfaces in bilingual language development: beyond structural overlap. *Int. J. Biling.* 13, 195–210. doi: 10.1177/1367006909339810
- Stojanovic, D. (1999). Parsing and acquisition: Evidence from Serbo-Croatian. Doctoral dissertation, University of Ottawa.
- Sun, C. (2006). *Chinese: A linguistic introduction*. Cambridge: Cambridge University Press.
- Sun, C., and Givón, T. (1985). On the so-called SOV word order in mandarin Chinese: a quantified text study and its implications. *Language* 61, 329–351. doi: 10.2307/414148
- Titov, E. (2017). The canonical order of Russian objects. *Linguist. Inquiry* 48, 427–457. doi: 10.1162/ling_a_00249
- Titov, E. (2020). Optionality of movement. *Syntax* 23, 347–374. doi: 10.1111/synt.12202
- Vasishth, S. (2002). *Working memory in sentence comprehension: Processing Hindi Center Embeddings*. Doctoral dissertation, Ohio State University.
- Vihman, V. A., and Nelson, D. (2019). Effects of animacy in grammar and cognition. *Open Linguistics* 5, 260–267. doi: 10.1515/opli-2019-0015
- Witzel, J., Witzel, N., and Nicol, J. (2012). Deeper than shallow: evidence for structure based parsing biases in second-language sentence processing. *Appl. Psycholinguist.* 33, 419–456. doi: 10.1017/S0142716411000427
- Yanko, T. E. (2001). *Kommunikativnye strategii russkoj rechi (in Russian, ‘communicative strategies in Russian speech’)*. Moscow: Yazyki slavyanskoj kul'tury.



OPEN ACCESS

EDITED BY

Stavros Skopeteas,
University of Göttingen, Germany

REVIEWED BY

Katalin Mády,
Hungarian Academy of Sciences, Hungary
Frank Kügler,
Goethe University Frankfurt, Germany
Corinna Langer,
Goethe-Universität, Germany,
in collaboration with reviewer FK

*CORRESPONDENCE

Janne Lorenzen
✉ janne.lorenzen@uni-koeln.de

RECEIVED 19 September 2023

ACCEPTED 25 March 2024

PUBLISHED 09 April 2024

CITATION

Lorenzen J, Roessig S and Baumann S (2024) Paradigmatic and syntagmatic effects of information status on prosodic prominence – evidence from an interactive web-based production experiment in German. *Front. Psychol.* 15:1296933. doi: 10.3389/fpsyg.2024.1296933

COPYRIGHT

© 2024 Lorenzen, Roessig and Baumann. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Paradigmatic and syntagmatic effects of information status on prosodic prominence – evidence from an interactive web-based production experiment in German

Janne Lorenzen^{1*}, Simon Roessig² and Stefan Baumann¹

¹IfL-Phonetik, University of Cologne, Cologne, Germany, ²Department of Language and Linguistic Science, University of York, York, United Kingdom

In this paper, we investigate how information status is encoded paradigmatically and syntagmatically via prosodic prominence in German. In addition, we consider individual variability in the production of prominence. To answer our research questions, we collected controlled yet ecologically valid speech by applying an innovative recording paradigm. Participants were asked to perform an interactive reading task in collaboration with an interlocutor remotely via video calls. Results indicate that information status is encoded paradigmatically via the F0 contour, while syntagmatic effects are subtle and depend on the acoustic parameter used. Individual speakers differ primarily in their strength of encoding and secondarily in the type of parameters employed. While the paradigmatic effects we observe are in line with previous findings, our syntagmatic findings support two contradictory ideas, a balancing effect and a radiating effect. Along with the findings at the individual level, this study thus allows for new insights regarding the redundant and relational nature of prosodic prominence.

KEYWORDS

prosody, information status, speech production, prominence, intonation, individual differences

1 Introduction

A crucial goal in communication is to signal discourse meaning via appropriate patterns of *relative prominence* among the words in an utterance. Prominence is a relational property that refers to a speech unit that “stands out” by virtue of a variety of factors pertaining to both meaning and form. Both paradigmatic and syntagmatic aspects have an influence on *prosodic* strength relations, and the aim of the present study is to investigate this interplay in German production data. We can think of the paradigmatic and syntagmatic aspects of prosodic prominence as two axes of the same concept. On the vertical or paradigmatic axis, we consider the prominence of entities occurring in the same phrasal position but encoding different discourse meanings. This axis takes into account prominence relations *across* different utterances. On the horizontal or syntagmatic axis, we consider the prominence of multiple successive entities. This axis hence accounts for prominence relations *within* a phrase or utterance. The layer of *meaning* we are looking at is the information status of referents, i.e.,

their level of givenness in discourse. The *form* relates to both phonetic and phonological features of prosodic analysis, namely gradual cues such as F0 height and excursion, duration or periodic energy, as well as as categorical distinctions between pitch accent type and status (i.e., prenuclear vs. nuclear).

In our formal analysis we follow the ‘metrical branch’ of prosodic phonology, proposed, for example, by Ladd (2008, 2014), and broadly defined as the hierarchical structure of utterances and their syntagmatic strength relations. A crucial insight is that there is no unified set of suprasegmental features and domains emerging from the metrical perspective (Ladd, 2014, pp. 50, 74). Nevertheless, there is a selection of phonological constituents which are considered relevant for metrical representations, organized in the *prosodic hierarchy* (e.g., syllable, foot, phonological word, intermediate phrase, intonational phrase, utterance). The tonal structure of the utterance adds another layer to this hierarchy. There are two types of tones associated with metrically important positions in the prosody hierarchy, which prototypically fulfill two different functions: boundary tones mark edges and are associated with higher-level phrases [e.g., H- indicates the end of an intermediate phrase (ip), H% marks the end of an intonational phrase (IP)], while pitch accents (starred elements, e.g., H*) are associated with prominent syllables and mark strong positions, or heads, in larger phrases. Figure 1A shows the abstract strength relation among prominent elements at the intermediate phrase level in a metrical tree, indicating that the nuclear accent is most prominent, and also structurally most important (only dominated by s-nodes; see, e.g., Calhoun, 2010). Pre- and postnuclear accents are secondary in relation to the nucleus, which is the only obligatory accent in an intermediate phrase. However, while the metrical tree adequately depicts the structural relations within the phrase (i.e., the postnuclear element is more closely tied to the nuclear element than prenuclear constituents would be) it does not mirror the actual prominence relation between pre- and postnuclear accents: the prenuclear element is only dominated by a w-node whereas the postnuclear element is dominated by an s-node at a higher level, although prenuclear accents are generally assumed to be more prominent than postnuclear accents (see Ladd, 2008, pp. 262–263; Calhoun, 2010, p. 3). In fact, this (empirically more reasonable)

relation can better be captured in a metrical grid, where prominence is assigned via the number of beats on units within an utterance (see Figure 1B, e.g., Hayes, 1995).

Although both paradigmatic effects of meaning-related factors and syntagmatic effects of (form-related) prosodic structure on prominence have been addressed independently in previous studies many times before, they have rarely been dealt with in conjunction. In this paper, we attempt to take a more comprehensive look at the (prosodic) prominence relations between two referents in a sentence and their influencing factors – broadened by a close investigation of speaker-specific differences.

1.1 Prosodic marking of information status in West Germanic (the paradigmatic perspective)

Metrical strength relations do not only depend on prosodic aspects of an utterance, but often reflect meaning-related choices, i.e., the semantic-pragmatic and syntactic properties of the utterance that are related to the previous discourse. For example, a difference in information structure (or rather its prosodic marking) can be expressed by a different mapping of information structural domains onto metrical structure (see Ladd, 2008; Calhoun, 2010), as in Figure 2. Here, the difference between the broad focus structure of the phrase *a cup of coffee* (Figure 2, left), which would be appropriate as an answer to the question ‘What would you like to drink?’, and the narrow (or contrastive) focus structure (Figure 2, right), being valid in a context such as ‘I’d rather like a pot of coffee and not...’, can be represented in an efficient and elegant fashion simply by reversing the weak and strong nodes. The minimalist tree is a shorthand for the fact that in broad focus, *coffee* occupies the strongest position (realized by the nuclear accent), whereas in narrow focus, the nuclear accent moves to *cup* (and *coffee* is deaccented). If more complex structures containing this phrase are built (e.g., *five francs seventy-five centimes and a cup of pretty tasteless coffee*; Ladd, 2008, p. 272), the w-s relation between *cup* and *coffee* still signals broad focus and s-w signals narrow focus. This representation of prosodic strength relations is arguably

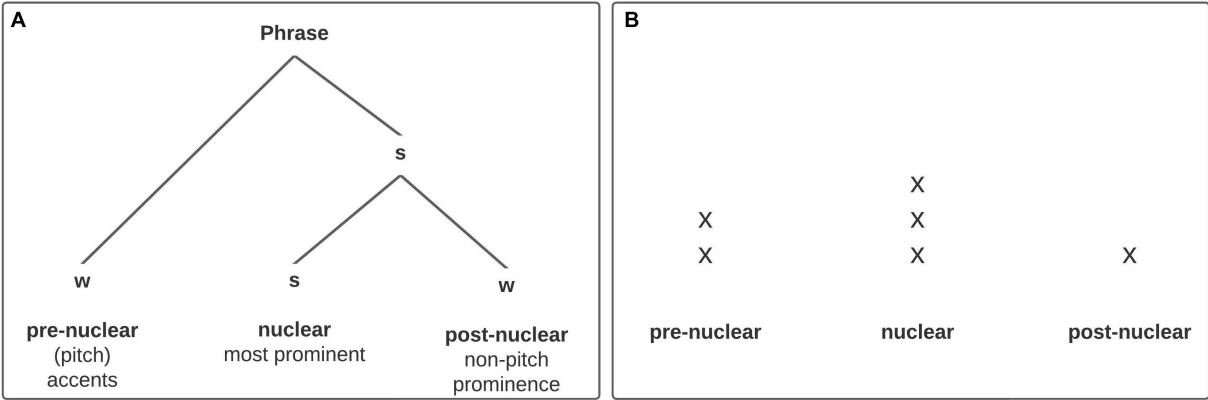
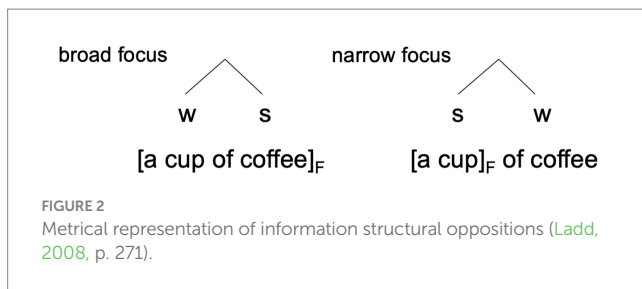


FIGURE 1 (A) Basic (ip-)phrase level metrical structure (Calhoun, 2010, p. 4); prenuclear and nuclear prominences are pitch accents, i.e., show tonal movement, while postnuclear prominences are non-tonal, especially durational. (B) The same structure presented as a metrical grid to illustrate the prominence relation between prenuclear and postnuclear accents more adequately.



very coarse-grained. There may be meaningful gradient variation in prosody, which is not captured in such models (e.g., Ladd, 2022; Roessig, 2024).

In the present study, we are concerned with information status, which can be regarded as another layer of information structure alongside focus. Following Chafe (1994), information status can be defined as the degree of cognitive activation or givenness of a referent in the discourse. It can be determined based on whether a referent has previously been established (often explicitly mentioned) in the current discourse and is thus already “active,” e.g., the second occurrence of *bus* in (1), or has been newly introduced from a previously “inactive” state, e.g., *a bus* in (1) and (2). Inactive referents are classified as *new*, active ones as *given*. A third category consists of *accessible* referents, which have not been explicitly mentioned but are retrievable through the context and thus are considered to be cognitively “semi-active,” e.g., *the driver* in (2).

- (1) I got on a bus (*new*) yesterday and the bus (*given*) was crowded.
(2) I got on a bus (*new*) yesterday and the driver (*accessible*) was drunk.

(examples adapted from Prince, 1981)

Information status has been linked to various, quite diverse concepts such as (shared) knowledge, consciousness, predictability as well as (un-)importance (see Prince, 1981; Baumann and Riester, 2012, for overviews). For the purpose of the present study, we will adopt Chafe’s idea of the information status of a referent or concept as part of the common ground between speaker and listener, which is more or less transparently derivable from the previous discourse context. For the sake of simplicity, we will restrict ourselves to Chafe’s three-way distinction of *new-accessible-given*, although more refined schematizations of information status are available, such as the *ReflEx scheme*, which distinguishes a lexical and a referential level of givenness (Baumann and Riester, 2012, 2013; Riester and Baumann, 2017). In our study, we focus on the contrast between *new* and *accessible* referents.

Different levels of information status are marked by prosody in German and other West Germanic languages, in particular via differences in prosodic prominence. Discourse-*given* referents are typically produced least prominently, discourse-*new* referents are prosodically most prominent and *accessible* referents occupy an intermediate position on the prosodic prominence scale (for German: Féry and Kügler, 2008; Baumann and Riester, 2013; for English: Ito et al., 2004; Chodroff and Cole, 2019; for Dutch: Swerts et al., 2002). Earlier accounts have focused on a binary distinction between *new* and *given* referents in terms of accentuation or deaccentuation, respectively (Halliday, 1967; Cruttenden, 2006; Ladd, 2008). However,

studies have shown that there is no one-to-one relation between deaccentuation and givenness but rather a probabilistic mapping (Terken and Hirschberg, 1994; Calhoun, 2010). Furthermore, like *new* referents, *accessible* referents are typically accented. Nevertheless, referents can be distinguished according to their information status via more fine-grained prosodic categories such as pitch accent type or gradient phonetic parameters such as F0 alignment and scaling, duration and intensity: *New* referents are typically produced with more high and rising accent types (e.g., Pierrehumbert and Hirschberg, 1990; Baumann and Riester, 2013) and with more extensive F0 excursion, later F0 peak alignment, longer duration and higher intensity when compared to *given* or *accessible* referents (e.g., Féry and Kügler, 2008).

Aylett and Turk (2004), drawing on information theory (cf. Shannon, 1948), proposed a principled account of the inverse relationship between ‘linguistic redundancy’ and ‘acoustic redundancy’ (i.e., prosodic prominence), known as the *Smooth Signal Redundancy Hypothesis* (SSRH). According to the SSRH, information is distributed evenly across the speech signal: Linguistically highly redundant information, which is already easily retrievable or predictable for the listener, is prosodically attenuated (i.e., acoustically less redundant or less prominent), while linguistically less redundant, harder-to-retrieve information is prosodically highlighted making it acoustically more salient (claimed at least for West Germanic languages). This yields a “smooth redundancy profile” in speech, ensuring robust communication, while at the same time reducing articulatory effort. The notion of predictability, which is central to the definition of linguistic redundancy, is directly related to information status, in that *given* referents can be regarded as highly predictable from the discourse context, *accessible* referents as somewhat predictable and *new* referents as unpredictable.

The original study conducted by Aylett and Turk (2004) confirmed this relation by showing that duration depends on word frequency, syllabic trigram probability and reference mention in English. This finding has since been extended, for example, to the relation between fundamental frequency modulations and predictability as captured by semantic focus and utterance probability (Turnbull, 2017). The Smooth Signal Redundancy Hypothesis thus systematically accounts for the paradigmatic relation observed between information status and prosodic prominence.

1.2 Prominence relations and the tonal context (the syntagmatic perspective)

It is not only the information status depending on the discourse context – and, as a consequence, the paradigmatic choices at the lexical and syntactic level – that influences an utterance’s metrical structure, but also the *syntagmatic prosodic context* within an utterance. It is well known, for example, that the prominence of a word or syllable can be enhanced by reducing the pitch of neighboring syllables (Gussenhoven et al., 1997), which is exploited, for example, in focus marking in the form of post-focal compression (e.g., Xu et al., 2004). Furthermore, the relative pitch height of two accents and their distance affect the perceived degree of prominence. Schettino and Wagner (2015) showed that in a German corpus of read speech, which has been annotated for prominence levels (Bonn Prosodic Database; Heuft, 1999), a pitch accent was judged as *more* prominent the further away it was from a previous accent. At the same time, speakers

produced these more distant accents proportionally somewhat less steeply and less delayed – and thus *less* prominently. This result not only confirms that we are dealing with both relative and cohesive patterns of prominence, but it is also in line with the declination effect (see [Pierrehumbert, 1979](#)), indicating that a lower peak accent following a higher one will be perceived as just as high and prominent as the preceding accent. This effect is even stronger the later the second accent occurs in an utterance.

Along similar lines, [Rump and Collier \(1996\)](#) investigated the effects of focus type on perceived appropriateness of pitch peak height in two referents within the same utterance. As expected, higher peaks were judged more appropriate by Dutch listeners for two subsequent referents that were both produced in contrastive focus than for two referents in broad focus (see [Figure 3](#)). Interestingly, however, when a single referent in contrastive focus was preceded or followed by a referent in the background, higher peaks were found to be more appropriate for the contrastive referent than when it was preceded or followed by another referent in contrastive focus. Similarly, [Bishop \(2017\)](#) found that in English SVO sentences, listeners dispreferred a prenuclear accent on the verb when the subsequent object was in narrow focus, but they showed no preference when it was in broad focus. These findings suggest that information structure is not only encoded locally (i.e., by position and type of the nuclear pitch accent) but that it is distributed across the sentence in the *prominence relation* of multiple accents.

In line with these observations, [Roessig \(2024\)](#) in a production study on German found an inverse relation of prosodic prominence between prenuclear and nuclear constituents depending on focus type. Prenuclear words occurring before nuclear words in corrective focus were realized less prosodically prominent (in terms of F0 height and syllable duration) than prenuclear words before words in (non-corrective) narrow focus and least prominently in a broad focus domain, while nuclear words in corrective focus were realized more prominently than nuclear words in narrow focus, which were in turn more prominent than nuclear words in broad focus. In another production study on German, [Kügler and Gollrad \(2015\)](#) also found that a prenuclear accent occurring before a nuclear accent in broad focus was realized with a higher peak F0 than a prenuclear accent occurring before a nuclear accent in contrastive focus. In a follow-up perception experiment they confirmed that listeners were more likely to interpret an utterance as conveying contrastive information when the F0 peak of the nuclear accent was higher than that of a prenuclear accent compared to when the F0 peak of the nuclear accent was lower than the prenuclear accent's. Information structure thus appears to be encoded in the *balancing* of the prenuclear and nuclear accents.

In contrast to these findings, [Gussenhoven and Rietveld \(1988\)](#) observed a positive correlation between the perceived prominence of two subsequent pitch peaks in Dutch listeners. In their prominence rating task, participants rated the prominence of a second peak lower if a preceding peak was realized with a lower F0 and thus less prominently, while a higher F0 on the first peak increased the prominence rating of a following peak. Higher prosodic prominence of the initial peak thus led to a perceived increase of prominence on a second peak irrespective of its realization. This suggests that prominence has a *radiating* effect, spreading onto following entities. This somewhat unexpected observation has become known as the *Gussenhoven-Rietveld Effect*. [Gussenhoven and Rietveld \(1988, p. 366\)](#) argue that the effect is potentially driven by the expectations of the listeners: If they encounter a low initial peak, they expect to hear another low peak next, which attenuates the perceived prominence of the actual peak that follows. At the same time, the study revealed the opposite effect for intensity, in that lower intensity of the first accented syllable increased the perceived prominence of the second one. Here, they argue that loudness is evaluated directly relative to the signal, i.e., lowering the intensity in one part of the signal leads to a perceived increase in loudness in the remaining parts. This observation seems to be more in line with the findings by [Rump and Collier \(1996\)](#) and [Roessig \(2024\)](#).

[Ladd et al. \(1994\)](#) partially replicated the Gussenhoven-Rietveld Effect with English listeners, but only in phonetically untrained participants and on a stimulus set containing a slightly lower second pitch peak (140 Hz). In a stimulus set containing a higher second peak (160 Hz), the opposite effect emerged in that an increase in the pitch of the first peak led to a decrease of perceived prominence on the second peak [akin to the balancing effects observed, for example, by [Rump and Collier \(1996\)](#) and [Roessig \(2024\)](#)]. This finding was confirmed in a second experiment with more values for the second pitch peak. According to [Ladd et al. \(1994\)](#), the Gussenhoven-Rietveld Effect replicated only in the condition with a lower pitch on the second peak since such a contour reflects what listeners consider a “normal” pitch range. A higher pitch range, on the other hand, implies an emphatic realization, which leads to the opposite effect. They concluded that while prominence is evaluated in a global fashion in non-emphatic productions, emphasis as a paralinguistic cue may override this interpretation ([Ladd et al., 1994, p. 98](#)).

To summarize, while there is evidence that prosodic prominence is influenced by the context at a syntagmatic level, it is not clear how such effects would materialize in the production of different levels of information status in German. In the case of a balancing or trade-off effect akin to the findings by [Rump and Collier \(1996\)](#) for the single contrastive focus conditions, [Kügler and Gollrad \(2015\)](#), [Bishop](#)

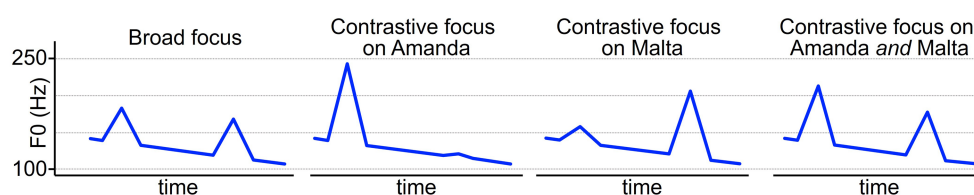


FIGURE 3

Target sentence and prototypical pitch contours for different focus conditions produced on the utterance *A'manda gaat naar 'Malta* "Amanda is going to Malta," adapted from [Rump and Collier \(1996, p. 9\)](#).

(2017), and Roessig (2024), an utterance could be expected to have a fixed prominence budget, which is distributed across the referents within an utterance according to their information status. For example, a *new* referent following or preceding an *accessible* referent would be more prominent than a *new* referent following or preceding another *new* referent. A radiating effect in production, analogous to the perceptual effect observed by Gussenhoven and Rietveld (1988) and Ladd et al. (1994), would imply that a more prominent referent in the utterance would also raise the prominence of the other referent, e.g., an *accessible* referent followed or preceded by a *new* referent would be more prominent than an *accessible* referent followed or preceded by another *accessible* referent. The contradictory findings in previous studies inspire different expectations for our present study, which calls for an exploratory approach.

1.3 Speaker-specific variability (the individual perspective)

Prosodic prominence is multidimensional, encompassing a variety of cues related to timing, changes in F0 and spectral characteristics of the speech signal (e.g., Fry, 1955; Sluijter and Van Heuven, 1996; Kochanski et al., 2005; Baumann and Winter, 2018; Roessig et al., 2022). Prosodically prominent entities are produced with steeper and more rising F0 contours, longer durations and higher intensity than less prominent entities in West Germanic languages. Due to the inherently redundant nature of prosodic prominence, pragmatic categories differentiated via prominence levels are thus encoded by a set of different cues producing the same effect. This redundancy may consequently enable a higher degree of individual variability, especially in the choice of prominence cues. Instead of encoding contrasts maximally redundantly by exploiting all prosodic cues to prominence, speakers may focus their production efforts only on specific cues.

Inter-individual variability is ubiquitous in speech production and perception and has recently received much attention in prosodic research. Previous studies have found, for example, considerable inter-individual variability in focus type marking. For instance, German speakers differed in the type and number of prosodic cues they employed to distinguish between broad, narrow and contrastive focus (Cangemi et al., 2015). Some speakers encoded a three-way contrast via multiple prosodic cues, e.g., peak alignment, peak height and word duration, while other speakers used only single cues to differentiate between two focus types, e.g., word duration to distinguish broad from narrow and contrastive focus and the presence or absence of prenuclear accents to distinguish broad and narrow focus from contrastive focus. These differences in the robustness of focus type encoding have implications for the interpretation of the utterances making some speakers more or less intelligible to the listener.

Similarly, Ouyang and Kaiser (2015) observed that American English-speaking individuals differed in how strongly they encode informativity (i.e., focus type, contextual probability, and word frequency) via the F0 contour, in that some speakers made larger differences in terms of F0 excursion between the levels of these three variables. Kim (2019) found considerable individual differences in prosodic cues to phrase boundaries in American English speakers. While all speakers produced pauses at IP boundaries, they differed in pause durations. In addition, most but not all speakers employed pitch

reset at IP boundaries and there was substantial variability in the scope of phrase-final lengthening across speakers.

Concomitantly, perception studies have observed substantial variability in the cues listeners attend to in the decoding of prosodic prominence (Cangemi et al., 2015; Baumann and Winter, 2018). Baumann and Winter (2018) in a study on German, for example, identified two groups of listeners, the larger group (about two thirds of the sample) attending primarily to cues related to the F0 contour, while the smaller group seemed to rely on non-prosodic cues such as word frequency, part-of-speech and argument structure in addition to duration. Cangemi et al. (2015) found different levels of proficiency in German listeners' ability to distinguish between focus types. What is more, specific listeners seemed to be particularly adept at interpreting focus types as produced by specific speakers but were less reliable in interpreting the productions of other speakers. In order to arrive at a comprehensive understanding of prominence production, we thus need to consider individual variability in our investigation.

1.4 Research questions and expectations

In this paper, we address three main research questions (RQs) related to paradigmatic, syntagmatic and individual aspects of the prosodic encoding of information status:

RQ1: How and to what extent is the information status of two successive referents encoded by prosodic prominence (at a paradigmatic level)?

RQ2: How and to what extent does the information status of two successive referents affect the prosodic prominence relation between them (at a syntagmatic level)?

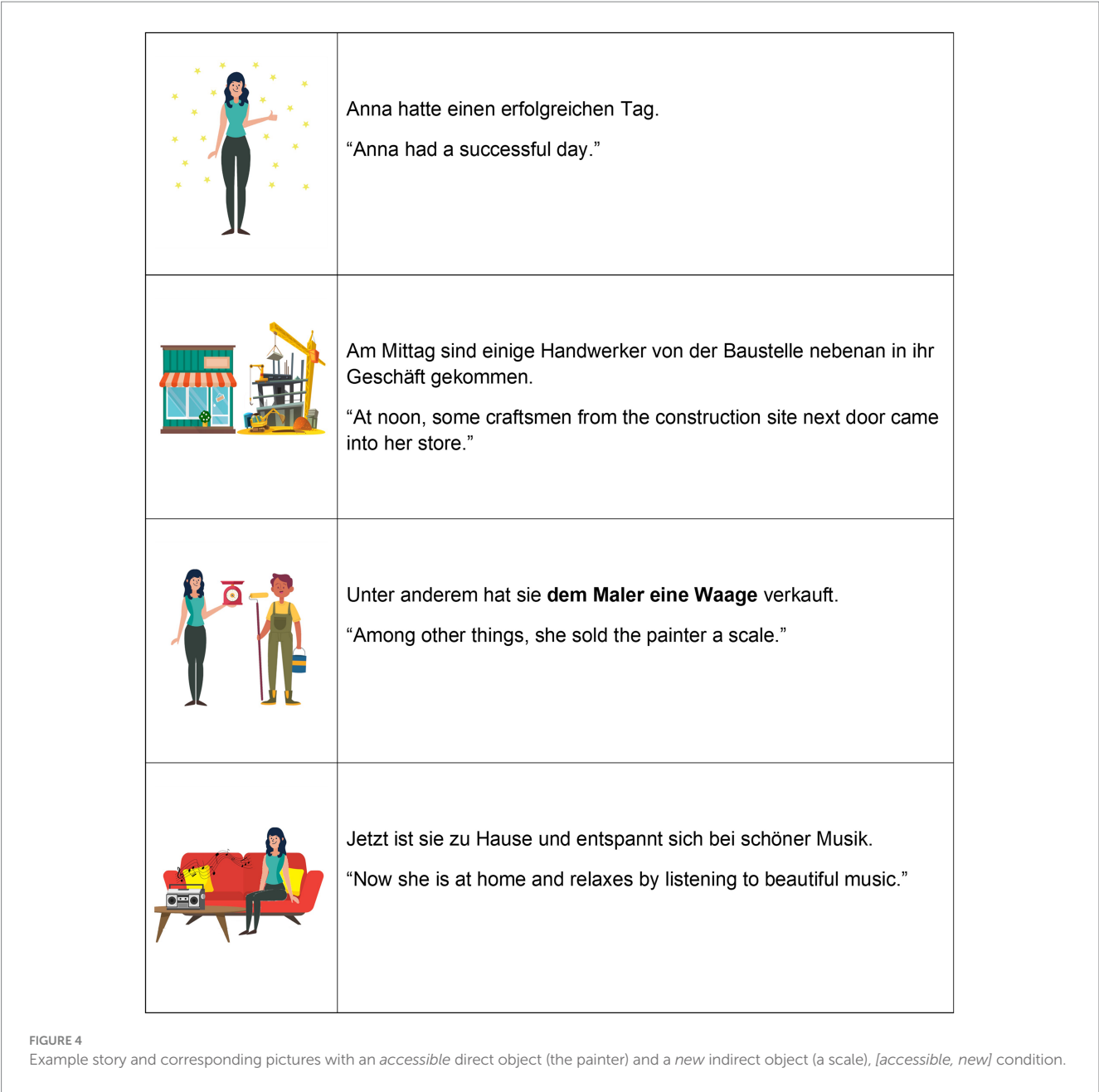
RQ3: Do individual speakers use different strategies in their encoding of information status and if so, how are these strategies characterized?

Based on previous research, we derive the following expectations:

(RQ1) Paradigmatically, *new* referents should be produced with higher prosodic prominence than *accessible* referents, both at the phonological and the phonetic level, e.g., by a larger number of rising accent types, with more extensive F0 changes, longer duration and followed by more phrase boundaries.

(RQ2) Previous studies suggest two potential outcomes concerning syntagmatic effects on prominence relations: a balancing effect or a radiating effect. Given that there seems to be more (recent) evidence in favor of the balancing effect, we assume this to be the more likely outcome, which we will thus take as a basis for the comparisons of the posterior estimates presented in Section 3.2.

(RQ3) We expect speakers to differ in the type of prosodic cues they employ as well as the strength of the encoding of the *new-accessible* contrast. Due to the morpho-syntactic marking of (in-) definiteness of each referent by the preceding article as a cue to information status, prosody may even be regarded as a redundant cue by some speakers.



2 Materials and methods

2.1 Reading material

To address our research questions, we collected data from a reading task. The analysis was based on 16 disyllabic target words with stress on the initial syllable, which were embedded in eight different short stories consisting of four sentences each (see Figure 4 for an example). Crucially, the target sentence, i.e., the third sentence in the story, included two consecutive target words in indirect and direct object roles. The indirect objects (Word1) always referred to people by their professions or, in one case, a family relation, e.g., *Lehrer* (“teacher”) or *Nonne* (“nun”). The direct objects (Word2) were everyday items, e.g., *Geige* (“violin”) or *Säge* (“saw,” see Appendix for full list of target words and stories). Target words belonged to comparable frequency classes according to the *Wortschatz Leipzig*

corpus.¹ We also ensured that the target words consisted of mostly sonorous segments to facilitate the analysis of the F0 contour.

The target words were either referentially and lexically *new* or *accessible* through the context provided in the previous sentence (in the example story in Figure 4, *the painter* is cognitively activated to some extent, i.e., *accessible*, by the scenario of a construction site mentioned before, while *a scale* is cognitively inactive, i.e., *new*).² *New* referents were preceded by an indefinite article, *accessible* referents by a definite article. We chose not to include *given* referents as they are often

1 https://corpora.uni-leipzig.de?corpusId=deu_news_2021
2 According to an annotation following the RefLex scheme (Baumann and Riester, 2012; Riester and Baumann, 2017), our *accessible* referents would be classified as *r-unused* (*r*=referential level) and *l-accessible* (*l*=lexical level). *New* target words would be annotated as *r-new* and *l-new*.

deaccented, which would pose difficulties for the comparisons of some parameters to accented referents. Each of the eight stories was devised in all four possible combinations of *new* and *accessible* target words ([*new, new*], [*new, accessible*], [*accessible, new*], [*accessible, accessible*]), but participants saw only one version of each story in a Latin square design. This served to prevent target words from becoming lexically *given* through repeated mentions. Each participant produced two stories for every possible information status combination. Target utterances were always produced in broad focus, either with a prenuclear or nuclear pitch accent on the target word.

2.2 Visual material

Speakers saw four pictures, each corresponding to one sentence of the story presented at the same time as the reading material. The pictures were also used in a staged picture sorting task (see Section 2.3 for more details). Pictures were created using resources from [Freepik.com](https://www.freepik.com).³ The first and last picture usually showed the protagonist of the story in different settings (in one case, the protagonist did not occur before the second picture, while the first picture showed a landscape). The second picture illustrated the setting of the main plot of the story (sometimes including the protagonist), which was related to one or both target words if they were *accessible* or unrelated if they were *new*. The third picture portrayed an interaction between the protagonist and another human interlocutor (the direct object, Word1) involving an item of interest (the indirect object, Word2). In this picture, both target words were displayed, which rendered them visually *given* for the speaker. The speaker was informed that the listener did not see the pictures until after the story was read aloud, so that the visual givenness was not common ground. As expected, we did not observe any deaccentuation of the target words, which would have been a strong indication of the referents being interpreted as *given* by the speaker. In addition, visual priming is known to only play a subordinate role in referent activation ([Baumann and Hadelich, 2003](#)).

2.3 Participants and data collection

We recorded 32 native speakers of German (8 identified as male, 24 as female, aged 20 to 38 years). Participants originated from seven different federal states of Germany and spoke no clearly detectable dialect. Two participants grew up bilingually with Russian or French as a second language but reported German to be their dominant language. Most were students at the time of recordings ($n = 27$). Participants were volunteers and received no compensation. They provided their written informed consent to participate in this study.

Recordings were collected remotely via a video call with the participant, the experimenter and a confederate in the summer of 2021, when COVID restrictions were still widely in place. To foster engagement

and thus prevent monotonous speech, participants were asked to perform an interactive task together with the confederate, whom they thought to be another participant of the experiment. The task was implemented as an animated browser app based on a Flask server with SocketIO written in Python. Participants read a short story aloud in such a way that the confederate would be able to memorize the story and, after a short delay, sort corresponding picture cards (see Section 2.2) into the correct order. While reading the stories, participants saw both the story and the corresponding pictures in the correct order (see [Figure 5](#), left panel). They were encouraged to first silently read and comprehend the story before reading it out loud. After reading, the participants were able to simultaneously watch the confederate sort the picture cards (see [Figure 5](#), right panel) and provide feedback on the correctness of the order of the pictures after the task was finished. During the picture sorting task, an additional two pictures were presented that were clearly unrelated to the story, which served as distractors to make the task seemingly more difficult for the listener.

The sorting was pre-programmed and always resulted in the correct order. Before the actual task started, there were two example stories to practice the procedure. In cases of hesitations, repairs, exaggerated segmental articulation, or continuation rises produced by the participant in the target sentences, they were asked to repeat the reading and sorting task of the affected story at the end with the experimenter feigning technical difficulties.

During the interaction, each participant wore headphones and sat in a quiet room at home in front of their computer, which served as their recording device. The recording itself was controlled by the experimenter via the podcasting app Ennuicastr.⁴ Ennuicastr records participants directly onto their own devices on separate audio tracks, which prevents unstable internet connections from distorting the audio. Zoom's built-in recording function was used as a back-up, which had to be used for 8 participants due to problems with the Ennuicastr recordings.

The use of Ennuicastr required minimal effort on the side of the participant. No prior installation of any software was needed since Ennuicastr runs directly in the browser. Participants joined the recording via a link that was created and sent to them by the experimenter. We recorded audio in lossless FLAC and downloaded the recordings in wav-format. However, the quality of recordings ultimately depended on the microphone quality on the side of the participant, which we were not able to control.

2.4 Data, annotation and measurements

Each participant read one version of each story, resulting in a dataset of 256 utterances including 512 target words. Recordings were force-aligned via WebMAUS ([Schiel, 1999](#); [Kisler et al., 2017](#)) and segment boundaries were subsequently manually corrected in Praat ([Boersma and Weenink, 2023](#)). The suprasegmental annotation was conducted independently by two trained annotators (one of them the first author of the present study) following the DIMA guidelines ([Kügler et al., 2022](#)). DIMA is an annotation system rooted in the autosegmental-metrical (AM)

³ Picture attributions for the pictures in [Figure 4](#): Living room, scale, hand - designed by Freepik; music notes - designed by brgfx/Freepik, painter, woman - designed by pikisuperstar/Freepik; radio, store, construction site - designed by macrovector/Freepik.

⁴ <https://ecastr.com/>

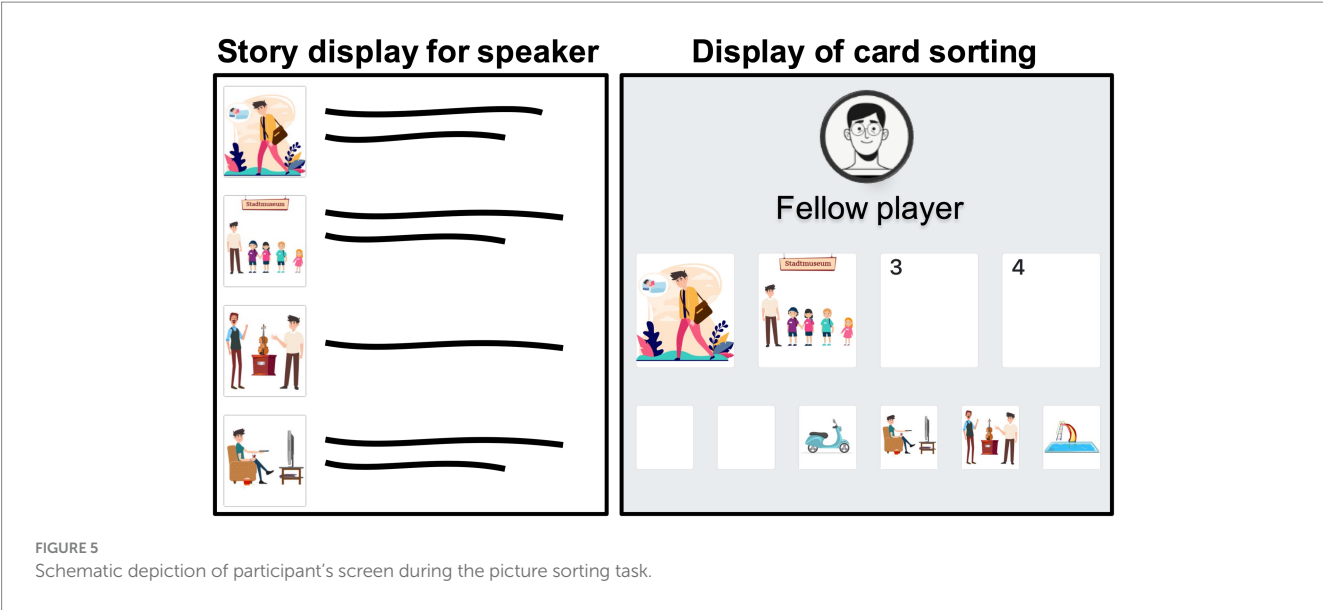


FIGURE 5
Schematic depiction of participant's screen during the picture sorting task.

TABLE 1 Summary of GToBI accent types and corresponding prominence scores following Baumann and Röhr (2015).

GToBI accent type	Accent type prominence score
L + H*	78.86
L* + H	71.53
H*	69.64
H + !H*	62.69
H + L*	57.14
!H*	53.62
L*	43.79

The scores represent mean ratings of perceptual prominence using a visual analogue scale ranging from 1 to 100.

framework of intonation analysis. It aims to be phonetically more transparent than other AM-based labeling systems (e.g., GToBI) thus facilitating annotation, but nevertheless reflects the phonological core of a contour. In cases of disagreement, a consensus was reached between the annotators and at least one other expert.

Specifically, two levels of phrase boundaries, strong (%) and weak (–), were annotated (agreement rate: 88%). Here, we consider the presence or absence of a phrase boundary after a target referent as a binary variable. Target utterances were frequently produced with a phrase boundary after Word1 (in 58% of cases) but rarely after Word2 (4%). On the tone tier, accentual and non-accentual tones were annotated, the latter as turning points occurring before and after pitch accents. For the accent type analysis, we translated the DIMA labels to GToBI (Grice et al., 2005) accent categories, for better comparability to previous studies (agreement rate: 89%).⁵ As a way to better compare

⁵ Note that these annotations are based on the DIMA tone labels, for which a consensus was already reached prior to the GToBI annotations.

the prominence of accent types across information status conditions that at the same time provides some perceptual validity, we assigned a prominence score to each accent type based on an independent prominence rating study by Baumann and Röhr (2015) with 68 native German listeners. In that study, participants rated the perceived degree of prominence of words on a visual analogue scale ranging from 1 to 100. The prominence scores reflect the average scores per accent type realized on the target word (summarized in Table 1).

Additionally, we measured several continuous phonetic parameters related to timing and the F0 contour. First, we measured the duration of the accented syllables of the target words in milliseconds. Target words that were audibly perturbed by internet connection issues (in cases in which we had to use the back-up Zoom signal) were excluded due to the duration measures being unreliable. In addition, phrase-final target words were not considered because of the potential effect of final lengthening. Furthermore, we included two measures from the ProPer toolbox (Albert et al., 2018; Albert, 2023), periodic energy mass and Delta F0. Both of these measures are based on periodic energy, which acoustically combines fundamental frequency and intensity by capturing the power of the periodic parts of the signal. This acoustic operationalization is motivated by perception, as periodic energy is correlated with pitch intelligibility (Albert, 2023, pp. 55–56).

Periodic energy mass quantifies the area under the periodic energy curve within an interval, typically the syllable (see Figure 6). It is the integral of duration and power. Since raw mass values are not deemed informative for various technical reasons (Albert, 2023, p. 150), relative mass is calculated by dividing the mass value for one syllable by the utterance's average mass value per syllable. The resulting unitless values are centered around 1, with values below 1 indicating weak mass and values above 1 indicating strong mass. Distorted recordings and phrase-final target words again had to be excluded, since the mass measure is contingent on duration and thus could also reflect effects of phrase-final lengthening. Since relative mass is determined using the utterance's average mass, we excluded both target words in an utterance that contained distorted audio.

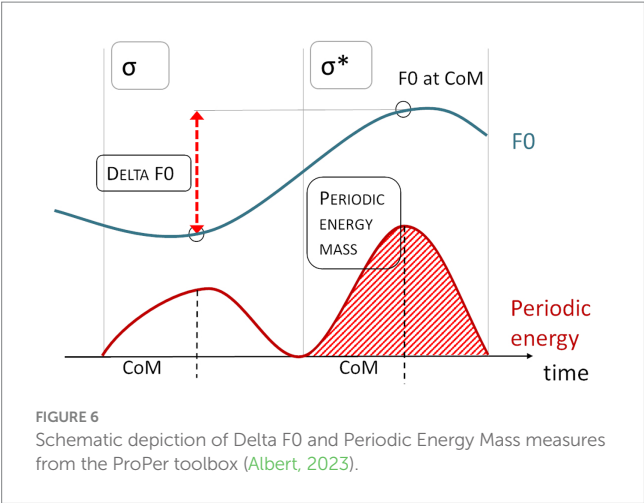


FIGURE 6
Schematic depiction of Delta F0 and Periodic Energy Mass measures from the ProPer toolbox (Albert, 2023).

TABLE 2 Summary of the parameters measured.

Parameter	Operationalization
Accent type prominence score	(Pseudo-)continuous, reflects perceived accent type prominence (in percent)
Delta F0	Continuous, F0 difference between accented and preceding syllable (in semitones)
Phrase boundaries	Binary, presence of boundary after target word
Syllable duration	Continuous, duration of accented syllables in non- phrase-final referents (in milliseconds)
Periodic energy mass	Continuous, integrates duration and power of the accented syllable (unitless)

Delta F0 captures the difference in semitones (st) between the F0 at the Center of Mass (CoM) of the accented syllable of a target word and the CoM of the syllable preceding it, thus reflecting F0 movement across syllables (see Figure 6). The CoM is the point in time that splits the area under the periodic energy curve of a syllable into two equal parts. Delta F0 is thus independent of landmark annotations and turning points, characterizing the F0 contour without the need for prior labeling. Table 2 summarizes all parameters considered here.

2.5 Statistics

We used Bayesian mixed effects linear regression models to investigate the paradigmatic and syntagmatic effects of information status on the prosodic realization of referents in two different positions. Note that the statistical analyses presented here are considered to have an exploratory character due to the complex nature of the research object and the relatively low token number.

We ran two models for each prosodic parameter, one to explore paradigmatic effects and one for the syntagmatic investigation. The prosodic parameter in question was always included as the response variable. The continuous phonetic parameters were z-scored. Phrase boundary was modeled as a binary variable.

For the investigation of the paradigmatic effects, each model included information status of the target word (levels: *new*, *accessible*) and position (levels: *Word1*, *Word2*) as well as their interaction as

predictors. The predictors were coded with treatment contrasts and the reference levels were *accessible* and *Word1*. In addition, we included random intercepts for word and speaker and by-speaker random slopes for information status and position to account for individual differences in the usage of the different acoustic parameters.

Random slopes capture the direction and size of an effect for each individual, which makes them ideal for the analysis of speaker-specific behavior. We extracted the by-speaker random slopes for information status and ran a hierarchical cluster analysis on these, following Baumann and Winter (2018) who performed a similar analysis using frequentist regression models. The cluster analysis served to group together speakers who follow similar strategies in their encoding of information status. Rather than considering each of the 32 individuals separately, which quickly becomes convoluted, the amount of the strategies that need to be considered is thus reduced in an objective manner.

To explore potential syntagmatic effects, we ran separate models, where we included information status as a variable with four levels, categorizing the information status of both referents in the utterance, i.e., [*new*, *new*], [*new*, *accessible*], [*accessible*, *new*], [*accessible*, *accessible*]. Again, position (levels: *Word1*, *Word2*) and the interaction of position and information status were included in the model. The reference levels for the treatment contrasts were [*accessible*, *accessible*] and *Word1*. We also included random intercepts for word and speaker and by-speaker random slopes for information status and position, similar to the paradigmatic models.

The models were fitted in R (R Core Team, 2022) using the Stan modeling language (Carpenter et al., 2017) via the *brms* package (Bürkner, 2017). Four sampling chains ran for 11,000 iterations each with a warm-up period of 5,500 iterations, yielding a total of 22,000 posterior samples per model. We used a weakly informative, normally distributed prior with a mean of zero and a standard deviation of ten for the regression coefficients and default priors supplied by *brms* for the remaining parameters. Model fits were assessed visually by inspecting the posterior predictive checks and by ensuring that no model yielded *Rhat* values larger than one. For each model, we report the regression coefficient β and 90% credible intervals (CIs) under the posterior distributions as well as the posterior probability that β is larger than zero: $Pr(\beta > 0)$. If zero is not included in the 90% CI and the posterior probability $Pr(\beta > 0)$ is larger than 0.95 or smaller than 0.05, we consider there to be compelling evidence that $\beta > 0$ or $\beta < 0$, respectively. Visually, the results of the models are depicted via half-eye plots generated using the *ggdist* package (Kay, 2023, see Schema in Figure 7). The colored density plots visualize the distribution of the posterior estimates, the thick and thin horizontal lines show 66 and 90% CIs, respectively, and the dot in the center of these lines represents the mean. The vertical dotted line marks the position where the estimate β equals zero, i.e., where there is no difference between the conditions that are being compared. When the thin horizontal black line does not cross the vertical zero line, the criterion that zero is not included in the 90% CI is fulfilled.

3 Results

3.1 Paradigmatic effects

We first present our findings on the paradigmatic effects of information status on the prosodic realization of referents. Target

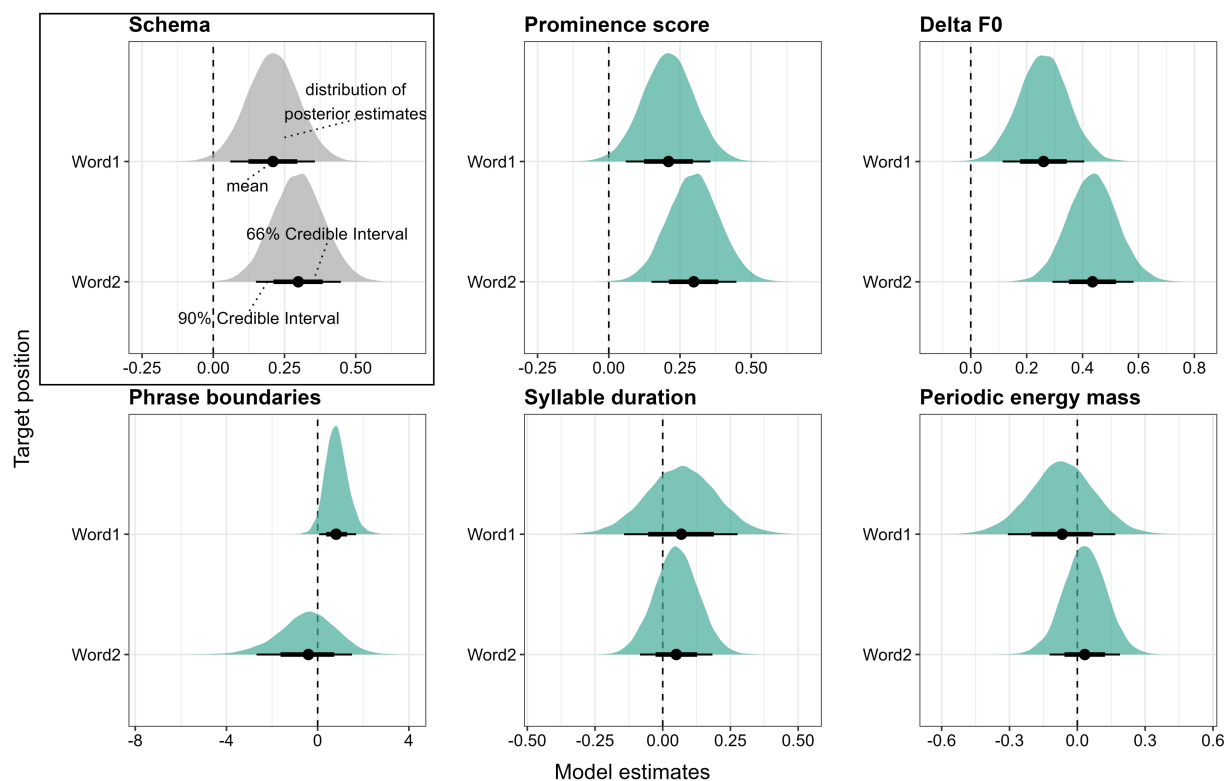


FIGURE 7
Posterior estimates for the five parameters as predicted by the models, means, 66% (thick horizontal lines) and 90% credible intervals (thin lines). Estimates show the change from *accessible* to *new* referents. Positive values indicate that the parameter is higher in *new* referents than in *accessible* ones.

utterances were always produced in broad focus with prenuclear or nuclear pitch accents on the target words. The bar plots in Figure 8 show proportions of GToBI accent types as a function of information status in both Word1 and Word2.⁶ In Word1, rising accents such as L+H* and L*+H are produced predominantly, while falling accents such as H+!H* and H+L* occur only in Word2. In Word2, H* accents are the most frequent accent type. Accent types in both positions often combine to form a “hat pattern” (e.g., Féry, 1993).

In *new* words, there is a larger proportion of L+H* accents than in *accessible* words in both positions. In Word1, L*+H, H* and !H* accents (the latter occurring only once) are used more frequently in *accessible* than in *new* words. In Word2, H* is more frequent in *new* words, but H+!H*, H+L* and !H* accents are more frequent in *accessible* than in *new* words. Following the ranking in Baumann and Röhr (2015), *new* referents are thus generally marked by more prominent accent types than *accessible* referents.

Converting the accent types to a numerical scale, i.e., the accent type prominence scores collected by Baumann and Röhr (2015; see Section 2.4), yields mean values of 73.2 (sd= 3.65) for *new* and 71.8 (sd= 4.24) for *accessible* referents in the first position, and mean values

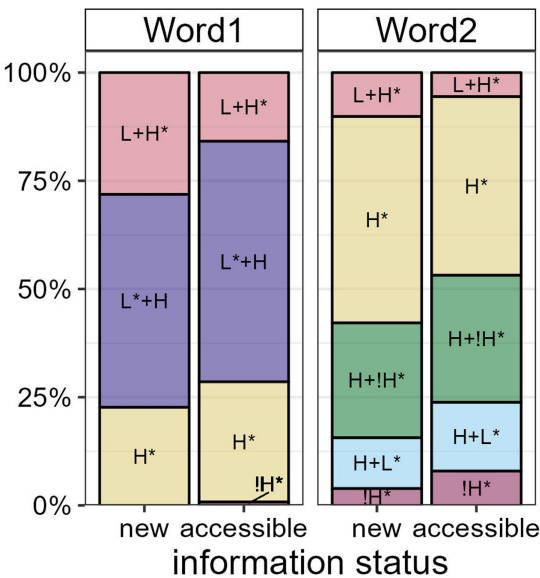


FIGURE 8
Proportions of GToBI accent types realized on Word1 and Word2 by information status.

6 Note that for better readability, we have collapsed L+^H* with L+H* and ^H* with H* accents. Furthermore, we excluded L* accents, which occurred only twice in the entire dataset.

of 66.6 (sd= 6.41) and 64.7 (sd= 6.68) for *new* and *accessible* referents, respectively, in the second position. The Bayesian mixed-effects model confirms that there is compelling evidence that *new* referents are

produced with higher accent type prominence scores in both Word1 ($\beta=0.21$, $CI=[0.06; 0.36]$, $Pr(\beta>0)=0.99$) and Word2 ($\beta=0.30$, $CI=[0.15; 0.45]$, $Pr(\beta>0)=1$). Note that while mean prominence scores are higher in Word1 due to the presence of many high rising accents in this position, the effect of information status is more robust in Word2 (see Figure 7, top center).

Next, we consider Delta F0. This variable captures two out of three tonal dimensions that have been shown to be relevant for perceived prominence in German (e.g., Baumann and Röhr, 2015): the direction of pitch movement and degree of pitch excursion but not the height of the accentual tone. Here, in line with Baumann and Röhr (2015), we assume that high Delta F0 (i.e., steeply rising pitch) is perceived as most prominent, while low Delta F0 (i.e., steeply falling pitch) is least prominent. There is evidence that the relation between these tonal aspects and prominence is more complex, for example, a steeply falling H+L* accent is perceived as more prominent than a slightly rising !H* accent. However, !H* is rare in our dataset so that we can assume a simplified relation between Delta F0 and prominence in our data.

Referents in Word2 exhibit on average lower Delta F0 than those in Word1, which is due to the frequently falling contour in this position. Crucially for our research question, *new* referents are produced on average with higher Delta F0 values than *accessible* referents. In Word1, the mean Delta F0 is 1.17 st (sd=1.99) in *new* target words and 0.60 st (sd=1.68) in *accessible* ones. In Word2, *new* referents are produced with an average Delta F0 of -0.40 st (sd=2.36) and *accessible* referents with -1.39 st (sd=2.03). As evident from Figure 7 (top right), these differences are confirmed to be reliable by the Bayesian model for Word1 ($\beta=0.26$, $CI=[0.11; 0.41]$, $Pr(\beta>0)=1$) and Word2 ($\beta=0.44$, $CI=[0.29; 0.58]$, $Pr(\beta>0)=1$). Similar to the results for accent type prominence, the effect of information status is more robust in Word2.

Next, we consider phrase boundary placement as the presence or absence of a boundary after a target word. Phrase boundaries are commonly placed after the first target word (in 58% of cases overall) but rarely after Word2 (only 4%). In the first position, phrase boundaries are more frequent after *new* target words (62.5%) than after *accessible* ones (53.5%). The Bayesian mixed-effects model confirms a reliable tendency for *new* referents in Word1 to be followed by more phrase boundaries than *accessible* referents ($\beta=0.83$, $CI=[0.07; 1.69]$, $Pr(\beta>0)=0.96$, see Figure 7, bottom left). The scarcity of boundaries after Word2 does not allow for strong conclusions and we will thus disregard this position.

In terms of syllable duration, accented syllables are on average shorter in Word1 (mean=223 ms, sd=52) than in Word2 (mean=254 ms, sd=51). Crucially, information status does not appear to have a systematic effect on duration (see Figure 7, bottom center). In Word1, accented syllables in *new* referents are on average 222 ms (sd=59 ms) long, and in *accessible* referents, they are 224 ms (sd=45 ms) long. In Word2, accented syllables are 256 ms (sd=49 ms) and 254 ms (sd=53 ms) long in *new* and *accessible* conditions, respectively. The durational differences between *new* and *accessible* referents cannot be considered reliable, neither in Word1 ($\beta=0.07$, $CI=[-0.14; 0.28]$, $Pr(\beta>0)=0.7$) nor in Word2 ($\beta=0.05$, $CI=[-0.08; 0.18]$, $Pr(\beta>0)=0.74$).⁷

Finally, most mass values are larger than 1, indicating relatively strong mass, which is unsurprising considering that we measure mass in accented syllables occurring in a stretch of speech containing mostly unaccented syllables. Mass values are on average larger in Word2 (1.85) than in Word1 (1.54). There is no systematic difference between mass values in *accessible* and *new* target words (see Figure 7, bottom right). In Word1, *new* referents are produced with an average mass of 1.51 (sd=0.48) and *accessible* referents with an average mass of 1.56 (sd=0.48). In Word2, both *new* and *accessible* referents exhibit an average mass of 1.85 (with standard deviations of 0.57 and 0.59, respectively). These differences do not prove to be reliable, neither in Word1 ($\beta=-0.07$, $CI=[-0.31; 0.17]$, $Pr(\beta>0)=0.32$) nor Word2 ($\beta=0.03$, $CI=[-0.12; 0.19]$, $Pr(\beta>0)=0.64$).

In summary, the overall results suggest that the paradigmatic contrast between *new* and *accessible* referents is reliably encoded via the F0 contour (i.e., GToBI accent type and Delta F0) and phrase boundary placement, but not via syllable duration or periodic energy mass.

3.2 Syntagmatic effects

Based on previous studies discussed in Section 1.2, we can assume that the production of prominence is also syntagmatically influenced by the context. That is, a referent may be more or less prominent depending on the prominence of a following or preceding referent. Here, we consider the information status of both referents in the utterance. Although this is an exploratory analysis and previous findings vary, we expect a balancing effect prior to our analysis, i.e., a *new* referent following or preceding an *accessible* referent should be more prominent than a *new* referent following or preceding another *new* referent (see Section 1.4). This expectation is formalized in the comparisons of the posterior estimates we conducted. Positive estimates will thus provide evidence in favor of the balancing effect, negative estimates support the radiating effect.

We investigate syntagmatic effects by keeping the information status of Word1 or Word2 constant and comparing the realizations of these target words depending on whether they are followed/preceded by a *new* or *accessible* referent. First, we focus on Word1. We compare the first target word in the [*accessible*, *accessible*] condition to the first target word in the [*accessible*, *new*] condition to determine how *accessible* referents in this position are produced depending on whether they are followed by an *accessible* or a *new* referent. In Figure 9 (top left), we can observe that an *accessible* referent followed by a *new* referent is produced with more prominent accent types than if it is followed by an *accessible* word. This tendency is in line with the radiating effect, but it is not confirmed to be reliable by our model ($\beta=-0.17$, $CI=[-0.37; 0.02]$, $Pr(\beta>0)=0.07$). Similarly, when comparing *new* referents in Word1 (i.e., the [*new*, *accessible*] and the [*new*, *new*] condition), there is evidence that those referents followed by other *new* referents are produced with more prominent accent types than those followed by *accessible* referents. This tendency is again not reliably predicted by the model ($\beta=-0.09$, $CI=[-0.3; 0.13]$, $Pr(\beta>0)=0.25$).

For Delta F0, results with regard to syntagmatic effects are even less conclusive (see Figure 9, top center). *Accessible* referents followed by *accessible* referents are produced with higher Delta F0 values than *accessible* referents followed by *new* referents, which is in the expected

⁷ Note that due to the exclusion of phrase-final referents, these findings are based on only 105 data points for Word1 and 238 for Word2.

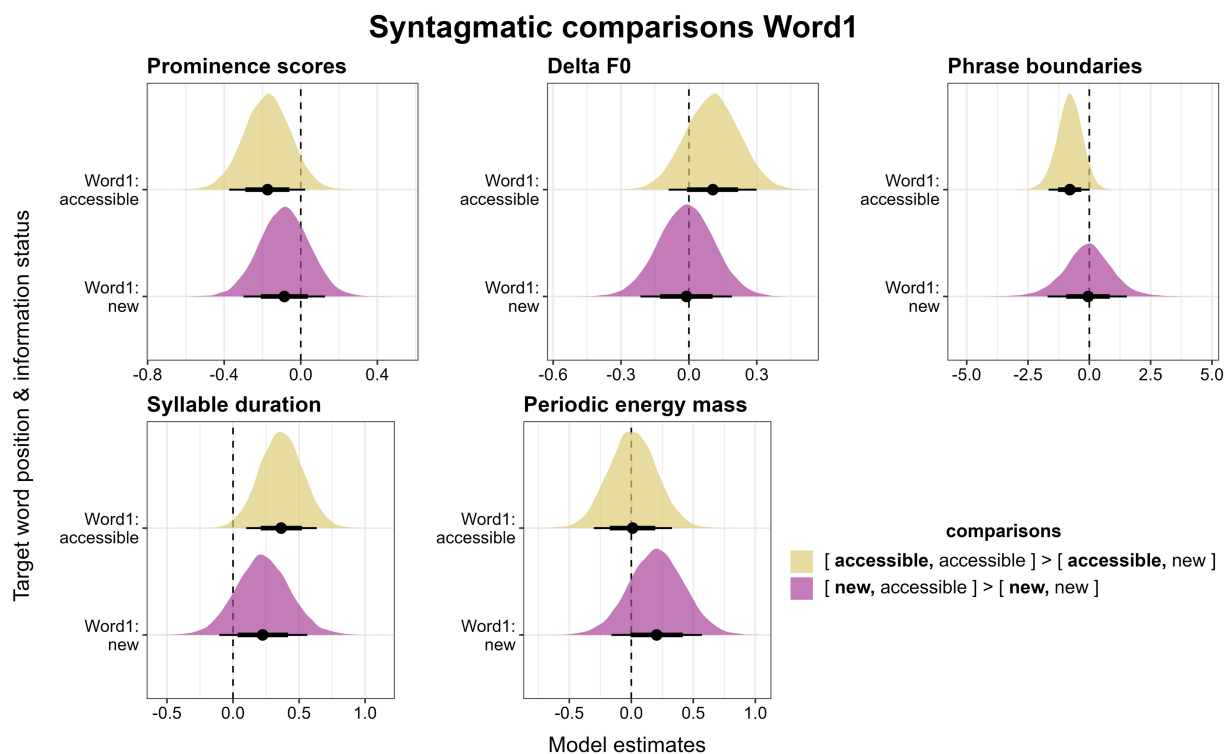


FIGURE 9

Posterior estimates for the five parameters as predicted by the models, means, 66% (thick horizontal lines) and 90% credible intervals (thin lines). Estimates show the changes for an *accessible* or a *new* referent in Word1 depending on whether it is followed by a *new* or an *accessible* referent. Positive values indicate that the parameter behaves according to our expectations (see "comparisons").

direction of the balancing effect, however, this difference is not reliable according to the model ($\beta = 0.1$, $CI = [-0.09; 0.3]$, $Pr(\beta > 0) = 0.81$). For *new* referents in Word1, there is barely a difference between those followed by other *new* and those followed by *accessible* referents ($\beta = -0.01$, $CI = [-0.21; 0.19]$, $Pr(\beta > 0) = 0.46$).

Next, we consider phrase boundaries, focusing only on the presence or absence of a boundary after Word1, since boundaries are rare after Word2. Boundaries are least often placed in the *[accessible, accessible]* condition (48%), but equally often after *new* referents irrespective of whether they are followed by an *accessible* or another *new* referent (both 63%). The trend for more boundaries in the *[accessible, new]* (59%) than the *[accessible, accessible]* condition is expected under the radiating effect if boundary placement is considered to boost the prominence of the preceding word. It seems fairly robust, but is not quite reliable according to the model ($\beta = -0.81$, $CI = [-1.66; 0.01]$, $Pr(\beta > 0) = 0.05$) (see Figure 9, top right). The difference between *new* referents followed by *accessible* or by *new* ones is clearly not reliable ($\beta = -0.06$, $CI = [-1.7; 1.53]$, $Pr(\beta > 0) = 0.48$).

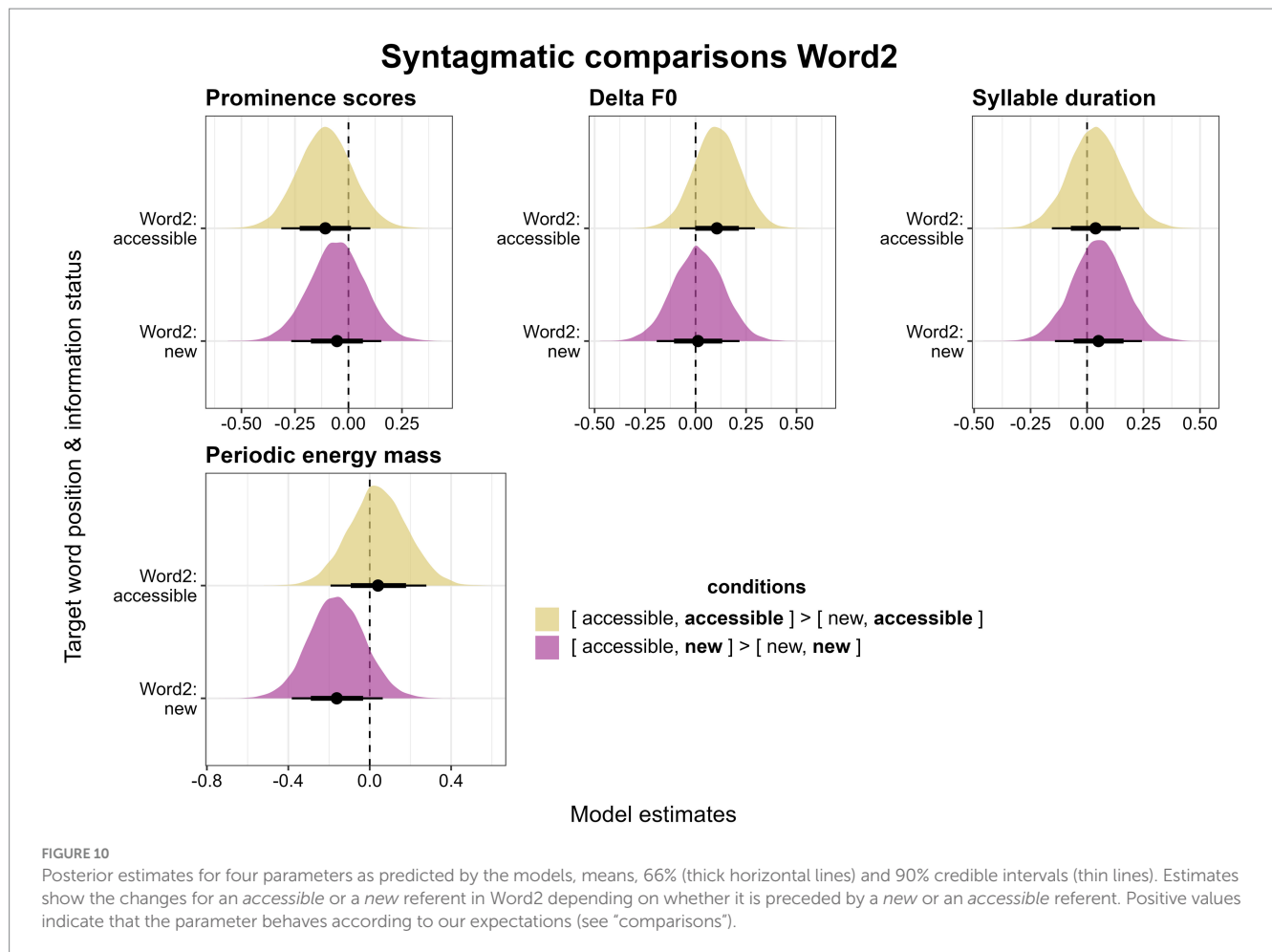
In terms of syllable duration, *accessible* referents in Word1 appear to be slightly longer when they are followed by another *accessible* referent than when followed by a *new* referent, which is in support of the balancing effect (Figure 9, bottom left). This difference is in fact reliable according to the model ($\beta = 0.37$, $CI = [0.1; 0.63]$, $Pr(\beta > 0) = 0.99$). However, we need to caution at this point that this finding is based on very few data points ($n = 32$ for *[accessible, accessible]* and $n = 25$ for *[accessible, new]*) due to the exclusion of phrase-final referents. For *new* referents, we observe the same trend in that they are longer when followed by *accessible* referents than when

followed by *new* referents. This difference, however, does not appear to be reliable according to the model ($\beta = 0.23$, $CI = [-0.1; 0.56]$, $Pr(\beta > 0) = 0.87$).

Finally, periodic energy mass does not exhibit any systematic syntagmatic effects (Figure 9, bottom right). In Word1, *accessible* referents barely differ based on whether they are followed by a *new* or an *accessible* referent ($\beta = 0.01$, $CI = [-0.3; 0.33]$, $Pr(\beta > 0) = 0.52$). *New* referents, on the other hand, seem to be slightly higher in mass when followed by an *accessible* referent than when followed by another *new* referent, which is in line with the balancing effect. However, this trend does not prove to be reliable ($\beta = 0.2$, $CI = [-0.16; 0.57]$, $Pr(\beta > 0) = 0.82$).

Turning now to Word2 (Figure 10), we keep the information status of target words in this position constant and compare two words with the same information status preceded by either a *new* or an *accessible* word. The same trends as for Word1 can be observed regarding accent type prominence scores (Figure 10, top left). Referents preceded by *new* referents are produced with more prominent accent types than those preceded by *accessible* referents. This is true for both *accessible* referents (comparing *[accessible, accessible]* to *[new, accessible]*) and for *new* referents (comparing *[accessible, new]* to *[new, new]*). Again, these tendencies support the radiating effect, but neither difference proves to be reliable (for *accessible* Word2: $\beta = -0.11$, $CI = [-0.31; 0.1]$, $Pr(\beta > 0) = 0.2$, for *new* Word2: $\beta = -0.05$, $CI = [-0.27; 0.15]$, $Pr(\beta > 0) = 0.33$).

For Delta F0, again similarly to Word1, *accessible* referents preceded by *accessible* referents (*[accessible, accessible]*) are produced with slightly higher Delta F0 values than *accessible* referents preceded by *new* referents (*[new, accessible]*, Figure 10, top center). This trend



goes in the expected direction of the balancing effect, however, it cannot be interpreted as reliable ($\beta=0.11$, $CI=[-0.08, 0.29]$, $Pr(\beta>0)=0.83$). Paralleling the findings for Word1, *new* referents in Word2 do not differ between the [accessible, new] and the [new, new] condition ($\beta=0.01$, $CI=[-0.19, 0.22]$, $Pr(\beta>0)=0.54$).

In Word2, there are no reliable differences in syllable duration between *accessible* referents preceded by *accessible* or *new* ones ($\beta=0.04$, $CI=[-0.16, 0.23]$, $Pr(\beta>0)=0.63$) nor between *new* referents preceded by *accessible* or *new* ones ($\beta=0.05$, $CI=[-0.14, 0.24]$, $Pr(\beta>0)=0.67$, Figure 10, top right). Finally, for periodic energy mass, there is again no difference between *accessible* referents preceded by *accessible* or *new* referents ($\beta=0.04$, $CI=[-0.19, 0.28]$, $Pr(\beta>0)=0.62$, Figure 10, bottom). However, *new* referents that are preceded by other *new* referents seem to have higher mass values than *new* referents preceded by *accessible* ones, which supports the radiating effect. This difference is not reliable ($\beta=-0.16$, $CI=[-0.38, 0.06]$, $Pr(\beta>0)=0.12$).

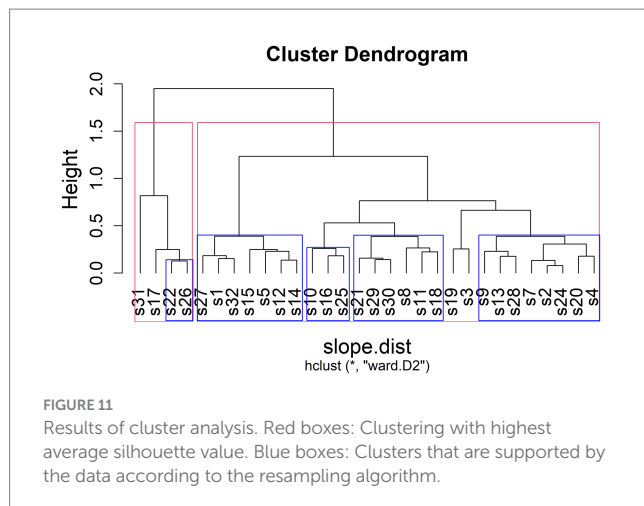
In summary, we find barely any strong evidence to support our *a priori* expectations that *new* referents should be produced more prominently when occurring before or after an *accessible* referent than before or after another *new* referent, while *accessible* referents should be less prominent when followed or preceded by a *new* referent than by another *accessible* referent (which we call *balancing effect*). Only syllable duration in *accessible* words in position Word1 exhibits a reliable difference in the expected direction. We also find some support for the opposite tendency (termed *radiating effect*),

which is weak, since most differences are not reliable according to our decision criteria, but still somewhat consistent: In all four information status combinations, *new* referents increase the prominence scores (derived from the pitch accent types used) of their neighboring referents, irrespective of whether they are *accessible* or *new*.

3.3 Individual strategies

Previous research has often discovered inter-individual variability in the encoding and decoding of prosodic prominence (see Section 1.3). Participants of these earlier studies often cluster into groups of speakers or listeners employing the same strategies. In order to identify such strategies in information status marking across speakers, we extracted the by-speaker random slopes for the effect of information status from four of the models presented in Section 3.1 (i.e., only the models with a continuous dependent variable).⁸ We ran a cluster analysis on these estimates. Following Levshina (2015), who

⁸ This analysis was conducted on a subset of 30 speakers, since for two speakers no slopes for syllable duration and periodic energy mass could be estimated due to the exclusion of data points.



suggests that average silhouette width serves as an indicator of the optimal number of clusters, a two-cluster solution is initially deemed the best fit for our data (see Figure 11, red rectangles). To validate the stability of the clusters, we used multiscale bootstrap resampling in the package *pvclust* (Suzuki et al., 2019). Results indicate that five sub-clusters (see Figure 11, blue rectangles) of our initial two clusters are supported by the data.⁹ In the following paragraphs, we will thus report on the characteristics of the five sub-clusters.

Figure 12 shows the averaged random slopes for each of the five clusters in comparison to the overall results. A positive value here indicates that this parameter is higher in *new* referents than in *accessible* ones and is thus employed to mark the information status contrast in the expected direction, a value around zero indicates that there is no difference and negative values indicate higher values in *accessible* referents as opposed to *new* ones.

Rather uniformly, the F0-related parameters emerge as the most important correlates of information status among all clusters. Periodic energy, however, is not used consistently in the expected direction by any cluster. With regard to duration, there is some variability across clusters. While in Cluster 5 (eight speakers), there is barely a difference in terms of syllable duration between *new* and *accessible* referents, Cluster 3 (three speakers) does use duration as a cue, at least to a larger extent than all other clusters. In Cluster 2 (seven speakers), syllable duration is as strong a cue as Delta F0 and accent type prominence score.

Clusters also differ in regard to the strength of encoding. While Cluster 1 (two speakers) subsumes particularly strong encoders, the remaining clusters distinguish between *new* and *accessible* referents much less clearly. Especially Cluster 2 seems to only weakly mark a difference in information status. Cluster 3 and Cluster 4 (six speakers) are very similar in their ranking of parameters, only differing in strength of encoding with Cluster 4 containing the weaker encoders.

It appears that the initial two-way split between speakers s22 and s26 (Cluster 1) plus two other speakers and the remaining speakers pertains primarily to the *strength* of encoding of the information

status contrast. Looking at smaller sub-clusters allows us to discover more fine-grained differences between groups of speakers and another dimension of variability, namely the *type* of parameters used to encode information status.

Individual speakers thus seem to differ among two dimensions, namely the strength of encoding of the *new-accessible* contrast and the choice of cues. However, most speakers make use of Delta F0 or accent type to encode information status.

4 Discussion

The innovative recording paradigm employed here enables us to investigate the prosodic encoding of information status in controlled yet less sterile material than usually elicited in experimental research. In addition, the paradigm provides an example of remote data collection in which participants are involved in an interactive task. Our results first and foremost confirm previous findings that information status is encoded via prosodic prominence in German, in that *new* referents are produced more prominently than *accessible* referents – which at the same time confirms our expectations regarding RQ1 on the paradigmatic perspective of the relation between information status and its prosodic marking. Our results are in line with the Smooth Signal Redundancy Hypothesis (Aylett and Turk, 2004; see Section 1.1), both at a phonological and a phonetic level. That is, *accessible* referents, which have become predictable through the discourse context, are produced with an attenuated acoustic signal, reflected not only by less prominent accent types and fewer insertions of prosodic boundaries after the target word but also by generally smaller F0 changes. *New* referents, on the other hand, are contextually unpredictable and thus produced here with relatively more prominent – and on average more steeply rising – accents, and are more often followed by a prosodic boundary, supposedly further enhancing the prominence of the preceding lexical item. Interestingly, other (continuous) measures such as duration or periodic energy, which have been shown to contribute to the encoding of prominence in various languages, were not modulated by the information status contrast. These findings further underline the precedence of F0 as a cue to phrase-level prominence (e.g., Baumann and Winter, 2018; Bishop et al., 2020; Roessig et al., 2022).

Turning to RQ2, we looked at the potential influence of the referents' information status on the prosodic prominence relation between them. As outlined in Section 1.2, syntagmatic effects are less well understood and different studies report conflicting findings. Our own syntagmatic findings are also less conclusive than our paradigmatic ones. A compelling previous finding that we expected to confirm prior to our analysis is that of a *balancing* or trade-off effect as, for example, observed by Rump and Collier (1996) or Roessig (2024). Here, prominence is distributed across referents in an utterance from a fixed budget: As the prominence of one referent increases, the prominence of another referent in the same utterance decreases. However, we observed this effect only in syllable duration and only reliably in one out of four comparisons. As for choice of pitch accent type and phrase boundary insertion, we find a different tendency, which was weaker yet more consistent: The presence of a prominent (i.e., *new*) referent in an utterance raises the prominence of another (preceding or following) referent in the same utterance, irrespective of its information status. This tendency is reminiscent of

⁹ Note that four speakers are not part of any of the five subclusters indicating that their behavior does not reliably cluster with any other speakers.

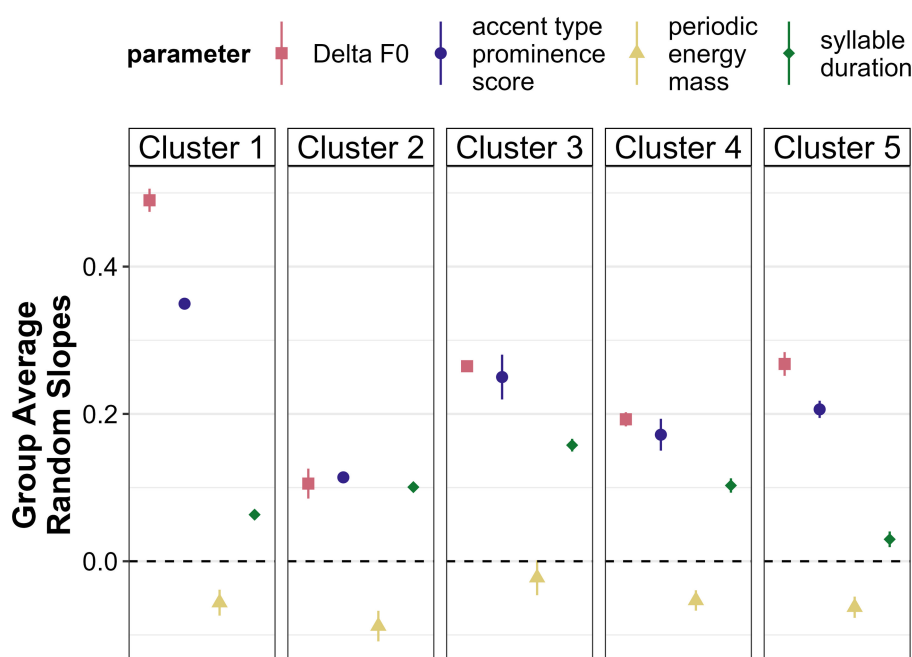


FIGURE 12

Mean random slopes and standard errors per cluster. The larger the value, the more strongly the cluster marks information status via this parameter in the expected direction.

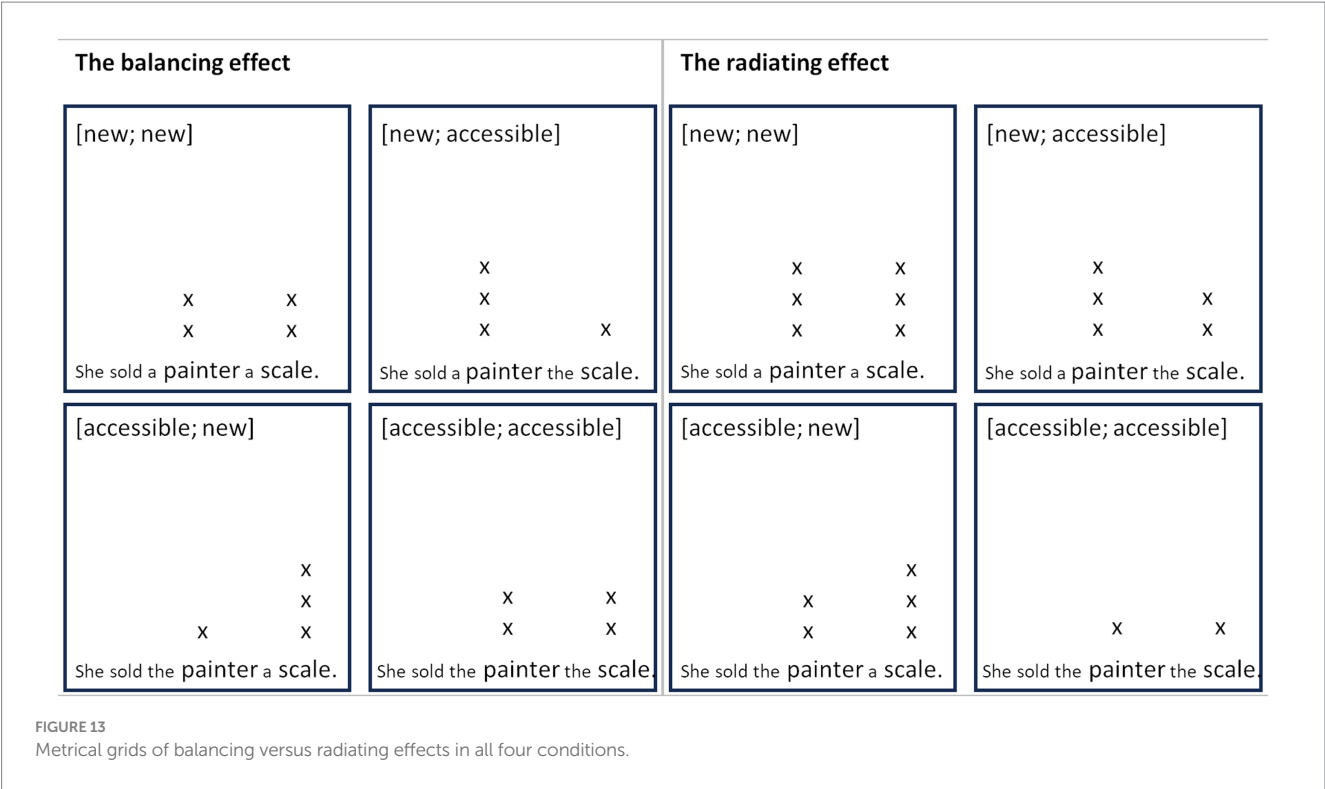
the Gussenhoven-Rietveld Effect, which describes that a more prominent pitch accent raises the perceived prominence of a following pitch accent (Gussenhoven and Rietveld, 1988). For the purpose of the present study, we refer to our comparable observation on the production side as a *radiating effect*, a term which is agnostic as to the direction of the process.

The contradicting findings of balancing versus radiating effects are systematized in the metrical grids in Figure 13. Metrical grids traditionally were not intended to capture subtle prosodic differences between pragmatic categories such as *new* and *accessible* referents (e.g., Liberman, 1975; Hayes, 1995). However, we make use of the flexibility they allow in the number of layers of beats to represent prominence relations between entities. Under the balancing effect, the number of beats assigned to all referents is equal in every single utterance (see Figure 13, left). The redistribution of prominence is indicated by the movement of a beat from a less prominent to a more prominent referent in the *[accessible, new]* or *[new, accessible]* condition as compared to the *[new, new]* or *[accessible, accessible]* condition, where the prominence relation between the two referents is balanced. Operating under the assumption of a fixed number of beats also implies that referents in *[new, new]* utterances are equally prominent as referents in *[accessible, accessible]* utterances. However, this is not what we observe in our data, as both referents in the *[new, new]* condition are realized more prominently than the referents in the *[accessible, accessible]* condition. Similarly, Rump and Collier (1996) find that both referents in double contrastive focus are more prominent than the referents in broad focus (see Figure 3 in Section 1.2). These observations are accounted for in the radiating effect by the addition of beats to more prominent referents (Figure 13, right). The presence of a *new*, i.e., a more prominent referent in the utterance raises the prominence of both referents in the utterance, so that a beat

is added to both referents in the *[new, accessible]* and the *[accessible, new]* conditions in comparison to the *[accessible, accessible]* condition. In order to keep the prominence relation between *new* and *accessible* referents tipped in favor of the *new* referent, another beat is added to the *new* referent in these conditions. In the *[new, new]* condition, the radiating effect raises the prominence and thus the number of beats in both referents simultaneously and to the same degree. That is, the radiating effect gives more weight to the paradigmatic influence of the respective semantic-pragmatic contrasts (in our case *new* vs. *accessible* information). In this respect, the radiating effect is not purely syntagmatic in nature, at least not to the same extent as the balancing effect is. In any case, both effects can be expected to occur in combination, and this is what we seem to find in our dataset as well.

The metrical grids represent an abstraction of the two effects observed in our data and in different previous studies. While the distribution of different layers of beats allows for some granularity in the representation of prominence as opposed to a binary branching tree (see Figure 2 in Section 1.1), one could ask whether the grids are fine-grained enough to capture the subtle changes in the realizations of referents we can attribute to either balancing or radiating effects in our data. The alternative is a continuous or gradient representation. This choice relates to a broader discussion in prosodic research on the categoricity and gradience of the suprasegmental signal (see, e.g., Grice et al., 2017; Roessig, 2021; Ladd, 2022).

In any case, the syntagmatic effects we observe are subtle. In fact, many contrasts do not reach the reliability threshold according to our statistical models. A potential explanation lies in the fact that most previous findings regarding syntagmatic effects of prominence, supporting either the balancing or the radiating effect, are based on perception experiments (Gussenhoven and Rietveld, 1988; Ladd et al.,



to produce this contrast prosodically. However, the use of indefinite articles for *new* referents and definite articles for *accessible* referents should have prompted the desired interpretation.

The robustness of the prosodic encoding of information status arguably has direct implications for perception: The more redundantly a speaker marks a given contrast in production, the more easily the listener can decode the message (Cangemi et al., 2015). However, it is unclear how the listener deals with contradictory prominence cues, e.g., when a speaker marks *new* referents as more prominent than *accessible* ones via the F0 contour, but at the same time *accessible* referents are higher in periodic energy than *new* ones (or vice versa). Another interesting question is whether modulating one parameter strongly or modulating several parameters moderately makes communication more robust. A corresponding perception study is needed to reveal the effectiveness of different prominence marking strategies.

5 Conclusion

The present study allows for new insights regarding the redundant and relational nature of prosodic prominence: (i) Paradigmatic effects prevail over syntagmatic ones in the encoding of information status. *New* referents are marked as more prominent than *accessible* referents, mainly by employing F0-related cues. (ii) In the context of the utterance, we observe both a balancing effect of prominence in terms of syllable duration and a radiating effect by which prominent entities appear to raise the prominence level of adjacent entities in terms of pitch accent type and phrase boundary placement. However, all observed effects are relatively small. (iii) There is substantial interspeaker variability, especially regarding the strength of the prosodic encoding of information status. The apparent redundancy of prominence marking often observed at the group level arises partly through different individuals prioritizing different prosodic cues.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://osf.io/frvm7/>.

Ethics statement

The studies involving humans were approved by Ethics Committee of the German Linguistic Society (Deutsche Gesellschaft für Sprachwissenschaft, DGfS). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

References

- Albert, A. (2023). *A model of sonority based on pitch intelligibility*, vol. 13. Berlin: Language Science Press.
- Albert, A., Cangemi, F., and Grice, M. (2018). Using periodic energy to enrich acoustic representations of pitch in speech: a demonstration. *Proceed. Speech Prosody* 2018, 804–808. doi: 10.21437/SpeechProsody.2018-162
- Aylett, M., and Turk, A. (2004). The smooth signal redundancy hypothesis: a functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Lang. Speech* 47, 31–56. doi: 10.1177/00238309040470010201
- Baumann, S., and Hadelich, K. (2003). Accent type and Givenness: an experiment with auditory and visual priming. *Proceedings of the 15th International Congress of Phonetic Sciences*, 1811–1814.
- Baumann, S., and Riester, A. (2012). “Referential and lexical givenness: semantic, prosodic and cognitive aspects” in *Prosody and meaning*. eds. G. Elordieta and P. Prieto (Berlin, New York: De Gruyter), 119–162.
- Baumann, S., and Riester, A. (2013). Coreference, lexical givenness and prosody in German. *Lingua* 136, 16–37. doi: 10.1016/j.lingua.2013.07.012

Author contributions

JL: Conceptualization, Data curation, Formal analysis, Writing – original draft, Writing – review & editing. SR: Conceptualization, Data curation, Formal analysis, Writing – review & editing. SB: Conceptualization, Data curation, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the German Research Foundation (DFG) as part of the SFB1252 Prominence in Language (Project-ID 281511265), projects A07 Metrical prominence – Scales and structures and A04 Dynamic modeling of prosodic prominence at the University of Cologne and project RO 6767/1-1 (Walter Benjamin program).

Acknowledgments

The authors thank Tiffany Zukas for creating the pictures used in the sorting task and her help with data annotation and Sölve Lorenzen for her help with data collection.

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2024.1296933/full#supplementary-material>

- Baumann, S., and Röhr, C. T. (2015). The perceptual prominence of pitch accent types in German. In *Proceedings of the 18th International Congress of Phonetic Sciences*, 298, 1–5.
- Baumann, S., and Winter, B. (2018). What makes a word prominent? Predicting untrained German listeners' perceptual judgments. *J. Phon.* 70, 20–38. doi: 10.1016/j.wocn.2018.05.004
- Bishop, J. (2016). Individual differences in top-down and bottom-up prominence perception. *Proceed. Int. Conf. Speech Prosody* 2016, 668–672. doi: 10.21437/speechprosody.2016-137
- Bishop, J. (2017). Focus projection and prenuclear accents: evidence from lexical processing. *Lang. Cogn. Neurosci.* 32, 236–253. doi: 10.1080/23273798.2016.1246745
- Bishop, J., Kuo, G., and Kim, B. (2020). Phonology, phonetics, and signal-extrinsic factors in the perception of prosodic prominence: evidence from rapid prosody transcription. *J. Phon.* 82:100977. doi: 10.1016/j.wocn.2020.100977
- Boersma, P., and Weenink, D. (2023). Praat: Doing phonetics by computer (6.3.10) [computer software]. Available at: [praat.org](https://www.praat.org)
- Bürkner, P.-C. (2017). Brms: an R package for Bayesian multilevel models using Stan. *J. Stat. Softw.* 80, 1–28. doi: 10.18637/jss.v080.i01
- Calhoun, S. (2010). The centrality of metrical structure in signaling information structure: a probabilistic perspective. *Language* 86, 1–42. doi: 10.1353/lan.0.0197
- Cangemi, F., Krüger, M., and Grice, M. (2015). “Listener-specific perception of speaker-specific productions in intonation” in *Individual differences in speech production and perception*. eds. S. Fuchs, D. Pape, C. Petrone and P. Perrier, vol. 3 (Frankfurt: Peter Lang)
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., et al. (2017). Stan: a probabilistic programming language. *J. Stat. Softw.* 76, 1–32. doi: 10.18637/jss.v076.i01
- Chafe, W. L. (1994). *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. Chicago: University of Chicago Press.
- Chodroff, E., and Cole, J. (2019). The phonological and phonetic encoding of information structure in American English nuclear accents. In *Proceedings of the 19th International Congress of Phonetic Sciences*, 1570–1574.
- Cruttenden, A. (2006). “The de-accenting of given information: a cognitive universal?” in *Eurotyp, 8, Pragmatic Organization of Discourse in the languages of Europe*. eds. G. Bernini and M. L. Schwartz (Berlin, Boston: De Gruyter), 311–356.
- Féry, C. (1993). *German Intonational patterns*. Tübingen: Niemeyer.
- Féry, C., and Kügler, F. (2008). Pitch accent scaling on given, new and focused constituents in German. *J. Phon.* 36, 680–703. doi: 10.1016/j.wocn.2008.05.001
- Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *J. Acoust. Soc. Am.* 27, 765–768. doi: 10.1121/1.1908022
- Grice, M., Baumann, S., and Benz Müller, R. (2005). “German intonation in autosegmental-metrical phonology” in *Prosodic typology: the phonology of intonation and phrasing*. ed. J. Sun-Ah (Oxford: Oxford University Press), 55–83.
- Grice, M., Ritter, S., Niemann, H., and Roettger, T. B. (2017). Integrating the discreteness and continuity of intonational categories. *J. Phon.* 64, 90–107. doi: 10.1016/j.wocn.2017.03.003
- Gussenhoven, C., Repp, B. H., Rietveld, A., Rump, H. H., and Terken, J. (1997). The perceptual prominence of fundamental frequency peaks. *J. Acoust. Soc. Am.* 102, 3009–3022. doi: 10.1121/1.420355
- Gussenhoven, C., and Rietveld, A. C. M. (1988). Fundamental frequency declination in Dutch: testing three hypotheses. *J. Phon.* 16, 355–369. doi: 10.1016/S0095-4470(19)30509-1
- Halliday, M. A. K. (1967). Notes on transitivity and theme in English: part 2. *J. Linguist.* 3, 199–244. doi: 10.1017/S0022226700016613
- Hayes, B. (1995). *Metrical stress theory—principles and case studies*. Chicago: University of Chicago Press.
- Heuft, B. (1999). *Eine prominenzbasierte Methode zur Prosodieanalyse und—Synthese*. Frankfurt: Peter Lang.
- Ito, K., Speer, S., and Beckman, M. E. (2004). Informational status and pitch accent distribution in spontaneous dialogues in English. *Proceed. Speech Prosody* 2004, 279–282. doi: 10.21437/SpeechProsody.2004-65
- Kay, M. (2023). Ggdist: Visualizations of distributions and uncertainty (v3.3.0) [Computer software]. Available at: <https://zenodo.org/>
- Kim, J. (2019). Individual differences in the production of prosodic boundaries in American English. In *Proceedings of the 19th International Congress of Phonetic Sciences*, 1024–1028.
- Kisler, T., Reichel, U., and Schiel, F. (2017). Multilingual processing of speech via web services. *Comput. Speech Lang.* 45, 326–347. doi: 10.1016/j.csl.2017.01.005
- Kochanski, G., Grabe, E., Coleman, J., and Rosner, B. (2005). Loudness predicts prominence: fundamental frequency lends little. *J. Acoust. Soc. Am.* 118, 1038–1054. doi: 10.1121/1.1923349
- Kügler, F., Baumann, S., and Röhr, C. T. (2022). “Deutsche Intonation, Modellierung und Annotation (DIMA) – Richtlinien zur prosodischen Annotation des Deutschen” in *Transkription und annotation gesprochener Sprache und multimodaler Interaktion*. eds. C. Schwarze and S. Grawunder (Tübingen: Narr), 23–54.
- Kügler, F., and Gollrad, A. (2015). Production and perception of contrast: the case of the rise-fall contour in German. *Front. Psychol.* 6:1254. doi: 10.3389/fpsyg.2015.01254
- Ladd, D. R. (2008). *Intonational phonology*. 2nd Edn. Cambridge: Cambridge University Press.
- Ladd, D. R. (2014). *Simultaneous structure in phonology*. 1st Edn. Oxford: Oxford University Press.
- Ladd, D. R. (2022). “The trouble with ToBI” in *Prosodic theory and practice*. eds. J. Barnes and S. Shattuck-Hufnagel (Cambridge: The MIT Press), 247–258.
- Ladd, D. R., Verhoeven, J., and Jacobst, K. (1994). Influence of adjacent pitch accents on each other's perceived prominence: two contradictory effects. *J. Phon.* 22, 87–99. doi: 10.1016/S0095-4470(19)30268-2
- Levshina, N. (2015). *How to do linguistics with R: Data exploration and statistical analysis*. Amsterdam, Philadelphia: John Benjamins Publishing Company.
- Liberman, M. (1975). The intonational system of English. Doctoral dissertation, Massachusetts Institute of Technology.
- Orrico, R., Gryllia, S., Kim, J., and Arvaniti, A. (2023). The influence of empathy and autistic-like traits in prominence perception. In R. Skarnitzl and J. Volin (Eds.), *Proceedings of the 20th international congress of phonetic sciences 1280–1284*. Guarant International.
- Ouyang, I. C., and Kaiser, E. (2015). “Individual differences in the prosodic encoding of informativity” in *Individual differences in speech production and perception*. eds. S. Fuchs, D. Pape, C. Petrone and P. Perrier, vol. 3 (Frankfurt: Peter Lang), 147–188.
- Pierrehumbert, J. B. (1979). The perception of fundamental frequency declination. *J. Acoust. Soc. Am.* 66, 363–369. doi: 10.1121/1.383670
- Pierrehumbert, J. B., and Hirschberg, J. (1990). “The meaning of Intonational contours in the interpretation of discourse” in *Intentions in communication*. eds. P. R. Cohen, J. Morgan and M. E. Pollack (Cambridge: The MIT Press), 271–311.
- Prince, E. F. (1981). “Toward a taxonomy of given-new information” in *Radical pragmatics*. ed. P. Cole, vol. 14 (New York: Academic Press), 223–255.
- R Core Team. (2022). R: A language and environment for statistical computing (4.1.3) [computer software]. R foundation for statistical computing. Available at: <https://www.R-project.org/>
- Riester, A., and Baumann, S. (2017). The RefLex scheme—annotation guidelines. SinSpec. Working Papers of the SFB 73. University of Stuttgart, 14. Available at: <http://elib.uni-stuttgart.de/bitstream/11682/9028/1/RefLex-SinSpec14.pdf>
- Roessig, S. (2021). *Categoriality and continuity in prosodic prominence*. Berlin: Language Science Press
- Roessig, S. (2024). The inverse relation of pre-nuclear and nuclear prominences in German. *Laboratory Phonol* 14, 1–40. doi: 10.16995/labphon.9993
- Roessig, S., Winter, B., and Mücke, D. (2022). Tracing the phonetic space of prosodic focus marking. *Front. Artificial Intelligence* 5:842546. doi: 10.3389/frai.2022.842546
- Rump, H. H., and Collier, R. (1996). Focus conditions and the prominence of pitch-accented syllables. *Lang. Speech* 39, 1–17. doi: 10.1177/002383099603900101
- Schettino, V., and Wagner, P. (2015). Prosodic prominence and contextual patterns: evidence from German. International conference on prominence in language, Cologne. Available at: <https://pub.uni-bielefeld.de/record/2734124>
- Schiel, F. (1999). Automatic phonetic transcription of non-prompted speech. In *Proceedings of the 14th international congress of phonetic sciences*, 607–610
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423. doi: 10.1002/j.1538-7305.1948.tb01338.x
- Sluijter, A. M. C., and Van Heuven, V. J. (1996). Spectral balance as an acoustic correlate of linguistic stress. *J. Acoust. Soc. Am.* 100, 2471–2485. doi: 10.1121/1.417955
- Suzuki, R., Terada, Y., and Shimodaira, H. (2019). Pvcust: Hierarchical clustering with P-values via multiscale bootstrap resampling (R package version 2.2-0) [computer software]. Available at: <https://CRAN.R-project.org/package=pvcust>
- Swerts, M., Krahmer, E., and Avesani, C. (2002). Prosodic marking of information status in Dutch and Italian: a comparative analysis. *J. Phon.* 30, 629–654. doi: 10.1006/jpho.2002.0178
- Terken, J., and Hirschberg, J. (1994). Deaccentuation of words representing ‘given’ information: effects of persistence of grammatical function and surface position. *Lang. Speech* 37, 125–145. doi: 10.1177/002383099403700202
- Turnbull, R. (2017). The role of predictability in intonational variability. *Lang. Speech* 60, 123–153. doi: 10.1177/0023830916647079
- Turnbull, R. (2019). Listener-oriented phonetic reduction and theory of mind. *Lang. Cogn. Neurosci.* 34, 747–768. doi: 10.1080/23273798.2019.1579349
- Winter, B. (2014). Spoken language achieves robustness and evolvability by exploiting degeneracy and neutrality. *BioEssays* 36, 960–967. doi: 10.1002/bies.201400028
- Xu, Y., Xu, C. X., and Sun, X. (2004). On the temporal domain of focus. *Proceed. Speech Prosody* 2004, 81–84. doi: 10.21437/SpeechProsody.2004-19



OPEN ACCESS

EDITED BY

Stavros Skopeteas,
University of Göttingen, Germany

REVIEWED BY

Yong-cheol Lee,
Hannam University, Republic of Korea
Valentina Saccone,
University of Florence, Italy

*CORRESPONDENCE

Emilie Destruel
✉ e-destruel@uiowa.edu
Aoju Chen
✉ aoju.chen@uu.nl

RECEIVED 22 December 2023

ACCEPTED 08 May 2024

PUBLISHED 25 July 2024

CITATION

Destruel E, Lalande L and Chen A (2024) The development of prosodic focus marking in French.
Front. Psychol. 15:1360308.
doi: 10.3389/fpsyg.2024.1360308

COPYRIGHT

© 2024 Destruel, Lalande and Chen. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

The development of prosodic focus marking in French

Emilie Destruel^{1*}, Louise Lalande² and Aoju Chen^{2*}

¹Department of French & Italian, University of Iowa, Iowa, IA, United States, ²Institute for Language Sciences, Utrecht University, Utrecht, Netherlands

Introduction: French is traditionally described as a language favoring syntactic means to mark focus, yet recent research shows that prosody is also used. We examine how French-speaking children use prosody to realize narrow focus and contrastive focus in the absence of syntactic means, compared to adults.

Method: We elicited SVO sentences using a virtual robot-mediated picture-matching task from monolingual French-speaking adults ($N = 11$), 4- to 5-year-olds ($N = 12$), and 7- to 8-year-olds ($N = 15$). These sentences were produced with narrow focus on either the subject or the object and contrastive focus on the object.

Results: Linear mixed-effects logistic regression modeling on duration, mean intensity, mean pitch, and pitch range of the subject and object nouns showed that the 4- to 5-year-olds did not use any of these prosodic cues for focus marking but the 7- to 8-year-olds distinguished narrow focus from non-focus through an increase in duration, mean intensity and to a lesser degree, mean pitch in the object nouns, largely similar to the adults, and tended to use mean pitch for this purpose in the subject nouns, different from the adults, who used duration.

Discussion: Our study corroborates previous findings that French-speaking 4- to 5-year-olds do not use prosody for focus. Further, it provides new evidence that 7- to 8-year-olds use prosody to mark narrow focus on the object in a more adult-like manner than narrow focus on the subject, arguably caused by a more dominant role of syntactic means in the subject position in French. Together, these findings show that syntax-dominance can influence both the route and the rate of acquisition of prosodic focus marking.

KEYWORDS

focus, prosody, acquisition, children, French

1 Introduction

The notion of *focus* plays a central role in communication by indicating to the hearer which part of the sentence is non-presupposed, and thus intended by the speaker to be asserted. Across languages, substantial variation is observed in the strategies used to signal focus. While some languages rely heavily on syntactic and lexical means, others primarily use prosodic means (Vallduví and Engdahl, 1996; Kügler and Calhoun, 2020). French is commonly described as favoring syntactic strategies, especially in marking focus on grammatical subjects, where clefting seems to be most appropriate (e.g., Lambrecht, 1994; Féry, 2001). Nevertheless, many studies on French prosody show that information structure does exert an influence on prosody. Post-focus deaccentuation, prosodic phrasing, phonetic cues, like pitch height and duration, and choice of tonal patterns all play a role in encoding focus when syntactic means are not readily available (e.g., Féry, 2001; Chen and Destruel 2010; Beyssade et al., 2015;

Lee, 2015). However, prosodic realization of focus in French-learning children has rarely been examined, in contrast to a wealth of research on children acquiring languages that primarily use prosody or use both prosody and word-order for focus-marking purposes (see Chen, 2018 for a review).

Our study seeks to address this gap by providing new empirical data on how French-speaking 4- to 8-year-olds use prosody to realize focus in French, compared to French-speaking adults. We are specifically interested in sentence-initial and final positions and two different types of focus, i.e., narrow vs. contrastive focus. Our study will further our understanding on how cross-linguistic differences in focus-marking shapes the rate and route of acquisition of prosodic focus-marking.

The remainder of the paper is structured as follows: We first briefly review literature on the realization of focus in French and French-speaking children's use of prosody in comprehending and realizing focus (section 2). We present our research questions and hypotheses in section 3, and discuss our methodology in section 4. We report on our statistical analysis and results in section 5. Finally, section 6 concludes.

2 Background

2.1 Focus and its realization in French

In the literature on information structure, the notion of *focus* is commonly understood as the part of the sentence that makes available a set of alternatives that the speaker takes to be salient, and which in turn conveys information about how utterances fit in to the larger discourse structure (Rooth, 1992; Krifka, 2008). For example, when focus is marked on the grammatical subject *Alice* (1a), it indicates that the alternative propositions relevant for interpretation are of the form “*x bought candy*.” Similarly, when focus is marked on the grammatical object *a pie* (1b), it indicates that the hearer knows Alice bought something but does not know what was bought.

- (1) a. [Alice]_F bought candy.
b. Alice bought [candy]_F.

Moreover, scholars have commonly distinguished (at least) two types of focus depending on how the focal alternatives are exploited in the sentence (Kiss, 1998; Gussenhoven, 2007; Zimmermann and Onea, 2011). First, the term *information focus* is used to refer to information that introduces completely new information into a discourse, often diagnosed by identifying the open variable in a congruent wh-question (2a). Second, a *contrastive focus* is taken as an alternative to an expression that has previously been introduced into the common ground, and often expresses a correction to false assumptions (2b).

- (2) a. Q: What did Alice buy? A: Alice bought candy (information focus).
b. Q: Did Alice buy popcorn? A: (No,) Alice bought candy (contrastive focus).

These two focus types can be accompanied by a variation in the scope of the focus domain: *broad focus* in (3a) refers to cases where entire syntactic constituents are focused (i.e., phrases, clauses,

sentences), whereas *narrow focus* in (3b) correlates to a single element (i.e., nouns, verbs, adjectives) (e.g., Gussenhoven, 2007).

- (3) a. Q: What happened? A: [Camille went to Paris]_F (broad focus).
b. Q: Who went to Paris? A: [Camille]_F went to Paris (narrow information focus).

While it is largely assumed that the focus element in a sentence is associated with prominence – compared to material that is already given in the discourse – the exact way to achieve this prominence may differ across and within languages (Kügler and Calhoun, 2020; Ladd and Arvaniti, 2023). To that end, West Germanic languages such as English and Dutch are known to primarily use prosodic means, i.e., placing a pitch accent on the focus element. Other languages such as Chadic languages rely on morphological markers. For example, Schwarz (2009) notes that in Buli, a morphosyntactic focus marker is required under the subject focus condition, which contains a particle either preceding or following the verb stem. This is illustrated in (4) below (from Schwarz, 2009: 952).

- (4) Who ate the beans? ‘THE/ THIS WOMAN ate them.’
nípōōwá lē ḡbōi.
woman.DEF PTL eat.

Complicating the matter further, a theoretical asymmetry between utterances with subject and object focus is also widely noted across different languages: the default position for prosodic prominence typically falls on the right-edge of a clause (i.e., in object position), leading subject focus sentences to be marked. French, the language of interest in this paper, illustrates this asymmetry well. Lambrecht (1994), among others, argues for a strong preference to use syntactic constructions like fronting and clefting (5) in the context of subject focus, while DestrUEL (2013) finds these are much less common in the context of object focus.

- (5) a. Camille, elle est. allée à Paris. ‘Camille, she went to Paris.’ (fronting)
b. C’est. Camille qui est. allée à Paris. ‘It is Camille who went to Paris.’ (clefting).

Accordingly, past studies on French prosody have argued that information structure can affect prosody when syntactic means are not available. For example, some research shows that the focused element can be marked by a specific rising contour, which is both higher in pitch and aligned later than a typical final accent in non-contrastive contexts (cf. *accent d’insistance*, Di Cristo, 1996, or focus accent, Jun and Fougeron, 2000). Other studies also find that the prefocal region is characterized by a reduced pitch range and amplitude of tonal movements and by a reduced number of phrase boundaries (Touati, 1987; Jun and Fougeron, 2000; Dohen and Loevenbruck, 2004), and that the postfocal region can be characterized by an absence of prominent pitch movements (Rossi, 1985; Touati, 1987; Di Cristo and Hirst, 1993; Di Cristo, 1996; Clech-Darbois et al., 1999; Jun and Fougeron, 2000; Féry, 2001; Dohen and Loevenbruck, 2004). More recent studies demonstrate that the initial rise, which is an optional tonal Low-High unit associated with the left edge of the accentual phrase (or AP, the smallest unit of phrasing above the

prosodic word level in French), can be an informative cue to focus in that it is more likely to occur at the left edge of a contrastive focus domain (D'Imperio et al., 2012; German and D'Imperio, 2016; Portes and Reyle, 2022). Yet, the clearest consensus across studies on focus and prosodic features in French is that prosodic phrasing, i.e., the grouping of words into phonological units of different sizes, is one of the main strategies employed by French speakers to mark the difference between given and new or contrastive elements. Specifically, a number of studies have reported a tendency for a narrow-focused constituent to be parsed in a separate AP (Féry, 2001; Dohen and Loevenbruck, 2004; Beyssade et al., 2009; Chen and Destruel, 2010; Michelas et al., 2014; Portes and Reyle, 2022). In operational terms, if two elements would typically be grouped into a single AP in a non-focused context, then in case one of them is focused, a prosodic restructuring process isolates that element in a separate AP.

In sum, French presents a great degree of variation due to prosody being only one of the options available to adult speakers in marking focus (Beyssade et al., 2009). This creates ambiguous input for children, which bears the question of how this might affect their acquisition. The next section turns to reviewing past literature on L1 acquisition of prosodic focus-marking in production across languages.

2.2 Acquisition of prosodic focus-marking in L1

Previous developmental research has generally suggested that although children can use prosody to distinguish new and given information to their interlocutor at the two-word stage in the second year of life, they do not acquire fully adult-like competence in production and comprehension of the prosody-to-information structure mapping until the age of 10 or 11 (Ito, 2018; Chen et al., 2020). Across languages, children acquire adult-like use of prosody in focus-marking in production at different rates and via different routes. Chen (2018) proposes that these differences are related to typological differences in prosodic focus-marking in the ambient languages, including reliance on phonetic means, transparency of form-function mappings, the use of the prosodic means in the lexical context and the importance of prosodic means in comparison to other means for focus marking. Specifically, children acquire the ability to use phonetic means, such as the phonetic implementation of phonological categories like lexical tones in Mandarin, lexical pitch accents in Swedish, and pitch accents in English, to distinguish narrow focus from non-focus, and differentiate different focus types at an earlier age in languages that exclusively rely on phonetic means for focus marking, like Mandarin (Yang and Chen, 2018). This happens compared to languages that use both phonological and phonetic means for focus marking, such as English (Hornby and Hass, 1970; Wonnacott and Watson, 2008), and Dutch (Chen, 2011; Romøren, 2016). Furthermore, children acquire phonological encoding of narrow focus earlier in languages with a more transparent form-function mapping between phonological means and focus conditions (e.g., Swedish vs. Dutch) (Romøren, 2016; Romøren and Chen, 2022). Transparency also affects phonological focus marking in different sentence positions within the same language. For example, Dutch-speaking children acquire phonological focus marking earlier in sentence-final position than in sentence-initial position where the phonological form-function mapping is less clear (Chen, 2011).

Moreover, the timing of acquiring pitch-related cues for focus marking differs based on whether pitch is also used for lexical purposes (e.g., Mandarin vs. Dutch), with children acquiring the use of duration cues earlier than pitch-related cues in languages like Mandarin. Lastly, the relative importance of prosody and non-prosodic means such as word order for focus marking affects children's use of phonetic means in distinguishing focus types in different syntactical settings. For example, 4- to 5-year-olds acquiring languages that use word order in conjunction with prosody to mark focus, like Finnish (Arnhold et al., 2016), use prosody more extensively and are less restricted by word order compared to children acquiring languages where prosody plays a primary role in focus marking, such as German (Sauermaun et al., 2011) and Dutch (Chen and Höhler, 2018).

With respect to French, the literature on children's acquisition of prosodic focus marking is very scarce compared to that on children acquiring a West Germanic language, and existing research has mainly examined perception and comprehension rather than children's production skills. More specifically, Szendrői et al. (2017) examined the acquisition of the comprehension of prosodic focus marking in English-, German- and French-speaking children but has relevant implications for French-speaking children's use of prosody in focus-marking in production and will be briefly reviewed here. The authors tested adults and children between the ages of 3 and 6 in their ability to recognize focal constituents in subject and object conditions through manipulations of prosody. All three languages allow the use of prosody in focus marking but differ in the use of syntactic means especially in subject focus (preferred in French, possible in German, dispreferred in English). In the experiment, the children were given visual stimuli in the form of three animal-tool pairs (subject-object, e.g., bird-hammer) and subsequently given false assumptions by the experimenter, where focus was marked on the target constituent depending on the condition. The children were expected to correct the target constituent (animal in subject condition and item in object condition) using the construction in (6), where the focal constituent is represented in italics.

- (6) a. Experimenter: The bird has the *bottle*, right? Child: No, the *hammer*. (object focus condition)
- b. Experimenter: The *bird* has the bottle, right? Child: No, the *hedgehog*. (subject focus condition)

Through this experiment, Szendrői et al. (2017) established that the English-, German- and French-speaking children could perceive focus in subject and object through prosodic cues and make corrections accordingly. However, they noted that the French-speaking children were more reluctant to give subject corrections, in comparison to the English-speaking children, suggesting that there are more natural ways to achieve this contrast in French, such as via clefting, and a possible difference in the acquisition of prosodic realization of focus in subjects and objects.

There is, to our knowledge, only two studies published on the production of focus by French-speaking children. First, Ménard et al. (2006) tested articulatory and acoustic correlates of contrastive focus in French-speaking 4- to 8-year-old children and adults, recording their repeated productions of the name 'Baba' in two carrier sentences as responses to the experimenter's questions in neutral and contrastive focus conditions. Results showed that the children were not adult-like in using variations in intensity, formant and articulatory strategies (i.e., lip movements) to signal contrastive focus by the age of 8. Second,

Esteve-Gibert et al. (2021) investigated the use of prosodic cues (i.e., syllable duration, word-level pitch range) and co-speech head gestures in the marking of focus types by French-speaking 4- to 5-year-olds, and the interaction between these two strategies. In their experiment, the participants were shown a visual display depicting a girl, Claire, with her eyes covered, and a bag containing one or more items. They were asked to interact with Claire by producing sentences containing noun phrases that would help her select a specific item in the bag indicated in the visual display (e.g., *prends [la moufle orange]_{NP}*—‘take [the orange mitten]_{NP}’). The trials with the bag containing just one item were used to elicit broad focus; the trials with the bag containing more than one item were used to elicit contrastive focus or corrective focus on either the noun or the colour adjective. Esteve-Gibert et al. (2021) found evidence for the use of gestures but no evidence for the use of syllable duration and pitch range expansion to distinguish focus types in the French-speaking 4- to 5-year-olds, different from their peers acquiring a language where prosody plays a larger role in focus marking, such as Finnish (Arnhold et al., 2016), German (Sauermaun et al., 2011) and Dutch (Chen and Höhler, 2018).

3 Research questions and hypotheses

Given the theoretical and experimental backdrop introduced in the previous sections, and the scarcity of studies on production of focus in French-speaking children, our study sets out to examine the following two research questions using a phonetic approach:

- i. How do French-speaking children aged 4 to 8 years modulate prosodic parameters to signal narrow information focus (hereafter narrow focus) and narrow contrastive focus (hereafter contrastive focus), compared to French-speaking adults?
- ii. Does the position of the focus constitute in the sentence (i.e., subject or sentence-initial vs. object or sentence-final) influence acquisition of prosodic focus marking?

Past studies on prosodic focus marking in French have been concerned with word- or syllable-duration (Lee, 2015; Esteve-Gibert et al., 2021), intensity (Lee, 2015), mean pitch (Lee, 2015), and pitch range (i.e., the difference between the highest and lowest pitch) (Esteve-Gibert et al., 2021; Portes and Reyle, 2022) in their phonetic analysis on the target words in different focus conditions. Following these studies, we focus on similar prosodic parameters at the word level, i.e., word duration, mean intensity, mean pitch, and pitch range.

Considering the findings from prior literature on acquisition of prosodic focus-marking and the role of prosodic and syntactic means in focus-marking in French, we hypothesize that

French-speaking children will be unable to manipulate the above-mentioned prosodic parameters to indicate focus, regardless of focus type and sentence position, at the age of 4- to 5-years, but will develop certain aspects of this ability by the age of 7- to 8-years (**Hypothesis I**). Indeed, prior literature suggests that, in languages that allow for prosodic focus marking but recognize other (non-prosodic) strategies as preferential, these preferred strategies may precede the dispreferred ones in the development of the given language (Szendrői et al., 2017). Consequently, speakers of French and other languages that prefer alternative (non-prosodic) strategies, such as syntactic means, may acquire prosodic focus marking later than speakers of languages that favour prosodic means, such as English, Dutch, Mandarin and Swedish. Further, due to the preference for other non-prosodic strategies for focus marking in subjects in French and in the light of Szendrői et al.’s (2017) finding on the asymmetry in French-speaking children’s comprehension of focus in subjects and objects, we hypothesize that French-speaking children will have earlier acquisition of prosodic focus marking in objects than in subjects (**Hypothesis II**).

4 Methods

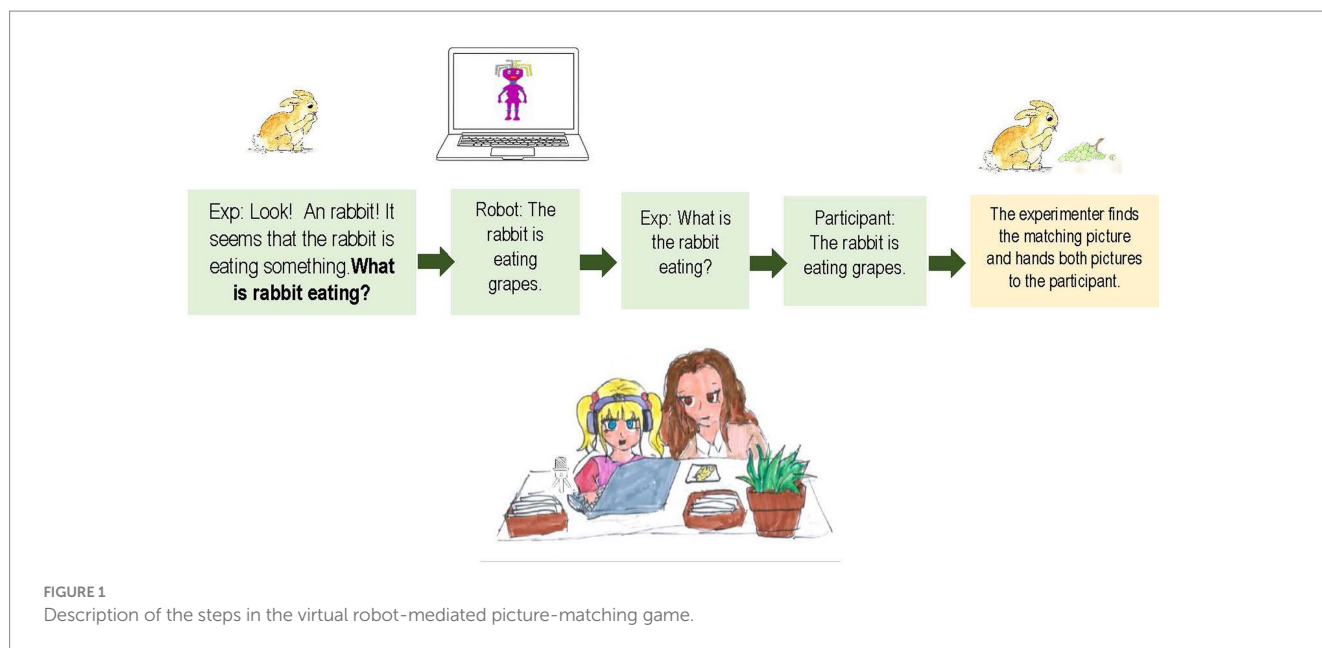
4.1 Participants

A total of 38 participants participated in this study. Of these, 11 were adult female speakers, 27 were children. All were monolingual native speakers of French, with normal hearing and vision. They all came from the Toulouse area in Southern France, and spoke French with no noticeable regional accent. The children were recruited from the Puygouzon primary school in the city of Albi. They consisted of two age groups: 12 were 4- to 5-year-olds and 17 7- to 8-year-olds. Details about their age range, average age and sex are provided in Table 1.

The two age groups were chosen for the following reasons. First, 4- to 5-year-olds are competent at producing multi-word sentences, which allowed us to study their use of prosody in focus marking in syntactically complete sentences. Second, previous studies on development in prosodic focus marking across languages (see Chen, 2018 for a review) have shown that children are adult-like in all or most of the aspects of prosodic focus marking at 7 or 8 years. Including 7- to 8-year-olds made it possible to study not only the developmental changes from 4- and 5-years to 7- and 8-years, but also find out whether French-speaking children become adult-like in prosodic focus marking at a similar age to children acquiring other languages. Finally, the eliciting task used in this study has been shown to be suitable for children aged 4 years and older. Using the same method to elicit data from two different age groups can facilitate the comparison in results between them.

TABLE 1 Participants’ biographical details.

	Age range (years;months)	Average Age (years;months)	Sex
Children group 1 (n = 12)	4.7–5.11	5.1	5 females
Children group 2 (n = 15)	7.2–8.3	7.1	8 females
Adults (n = 11)	17–44	31.4	11 females



4.2 Materials and procedure

The elicitation task in this study was adapted from the picture-matching game used by [Chen \(2011\)](#), which was also used in previous studies on children acquiring typologically different languages ([Chen, 2018](#)). It was designed to elicit SVO declarative sentences in the form of responses, where focus was marked in varying positions according to the target condition. Since the subject NP was always in sentence-initial position and the object NP in sentence-final position, we refer to the locus of focus as the position of focus and use the terms initial-position and final position interchangeably with subject and object in what follows. We tested three focus conditions: (i) narrow focus on the subject (NSF) which was triggered by who-questions (7a), (ii) narrow focus on the object (NOF), triggered by what-questions (7b), and finally (iii) contrastive focus on the object (NOCF), formulated as corrections to false assumptions (7c).

- (7) a. Q: Qui mange les raisins? ‘Who eats the grapes?’
A: [Le lapin]_F mange les raisins. ‘[The rabbit]_F eats the grapes.’
- b. Q: Que découpe la mamie? ‘What does the grandma cut?’
A: La mamie découpe [le jambon]_F ‘The grandma cuts [the ham]_F.’
- c. Q: Est-ce que le garçon peint une fleur? ‘Does the boy paint a flower?’
A: Non, le garçon peint [un ballon]_F ‘No, the boy is painting [a ball]_F.’

The procedure took the form of a structured game, in which the participant and the experimenter sat in front of a computer screen and boxes of pictures together and the participant was supposed to help the experiment with finding the matching picture for each of her pictures. On each trial, the experimenter

took an incomplete picture from a box (e.g., a rabbit that seems to be eating something). The experimenter then drew the participant’s attention to the picture and established what the picture was by stating “Look, a rabbit!” The experimenter then described what seemed to be missing in the picture and asked the congruent wh-question about the missing entity in the NSF and NOF conditions, which could be either the subject or the object (see [Figure 1](#)). In the NOCF condition, the experiment asked a yes-no question, making a guess about the missing object (e.g., *Is the rabbit eating a strawberry?*).

Upon hearing the question on each trial, the participant turned to a virtual robot assistant on the computer screen, which was there to help them get the correct answer. The participant clicked on the robot picture displayed on the screen and heard the answer in SVO form that included the missing part, via their headphones. The robot’s sentences were constructed from words recorded by a female native of French (age 35, speaker of standard Parisian French) in a randomized word list such that they contained no sentence-level prosody. Then, the experimenter repeated the question and the participant was instructed to respond to the question, using the exact same words used by the robot but speaking normally, instead of sounding like the robot.¹ Finally, the experimenter looked for the picture of the missing part in the box and handed both pictures to the participant.

A total of 28 pairs of images were created, spread over 4 practice items, 8 NSF, 8 NOF and 8 NOCF. The subject and object nouns were all disyllabic words that were familiar to French-speaking 4- to 5-year-olds. These words were taken from French textbooks used in elementary schools. Where possible, words with sonorants and voiced

1 Although the more natural way to respond to an open question is with a single constituent answer, sentences using the canonical order Subject-Verb-Object are commonly found in French, even in the spontaneous speech of 4-to-8-year-old children.

consonants were used to facilitate annotation for pitch analysis. Each subject noun in the NSF condition occurred also in the NOF condition; each object noun occurred in all focus conditions. A full list of stimuli is provided in the [Supplementary material](#) (section 1.2).

The adults were tested individually in a quiet room by a female native speaker of French (first author) on the campus of the university of Albi, and were paid a small fee for their participation. The children were tested individually by a near-native French-speaking female research assistant on site in their classroom, but outside of class time. Each participant was randomly assigned to one of two semi-randomised stimulus orders. One stimulus order was the reversed order of the other. Each session began with a brief chat to help the participant become more comfortable with the experimenter. Following this, the experimenter explained the game and started the game with four practice trials. If a child provided elided or full-sentence answers but with pronouns or appeared to imitate the robot's way of speaking during practice, the experimenter reminded him/her of the game rules. This was only needed for a small number of 4- to 5-year-olds. Most children responded with full sentences in their own prosody right from the start. Each test sessions lasted approximately 45 minutes for the younger children, about 35 minutes for the older children and less than 30 minutes for the adults. All sessions were audio-recorded using a portable digital recorder with a 48kHz sampling rate, 16-bit resolution, and an external high-quality microphone positioned 10–15 cm from the participants. Nevertheless, the participants were not instructed to stay seated in the exact same position, so the distance between the microphone and the children's mouth could not be precisely controlled for.

4.3 Data selection and annotation

The audio recordings from each participant were first orthographically annotated in Praat ([Boersma and Weenink, 2022](#)). During that first phase of the analysis, we identified three types of issues in children's production, which we labeled rephrasing, fluency and experimenter errors. *Rephrasing* describes trials in which participants did not exactly answer the questions using the robot's words, but instead used alternative focus strategies like clefting or fronting, inserted additional words (e.g. *non* 'no' preceding the answer in the contrastive condition), or gave a one-word answer. *Fluency* refers to trials in which participants hesitated (identifiable by long pauses), copied the robotic speech patterns, or used inadequate speech, such as speaking too rapidly or too softly. Finally, *experimenter error* refers to the trials in which the question was not repeated before a given response. The trials produced with one of these errors were removed from analysis.

Moreover, if the same errors were frequently produced by the same participant across stimulus items in a specific condition (or throughout all conditions), the data from that condition or that participant altogether was removed from analysis. This decision led to removing data from two participants in the younger group (4–5 years olds). Additionally, one of the items in the NOCF condition was eliminated in the analysis of all participants due to an issue in the formulation of the answer given the question asked.²

Thus, instead of 8 items, we report only 7 in the NOCF condition. The [Supplementary material](#) (section 1.1) gives an overview of the number of usable trials, per condition and group, that entered our analysis.

The target words ($n = 1,113$), i.e., the subject and object nouns, in the usable sentences ($n = 575$) were then annotated for word boundaries, following standard procedures ([Machač and Skarnitzl, 2009](#)) using Praat ([Boersma and Weenink, 2022](#)). Acoustic measurements including duration (ms), mean intensity (dB), mean pitch and pitch range (Hz) were subsequently extracted from these words using ProsodyPro ([Xu, 2013](#)). Finally, mean pitch and pitch range were checked for possible pitch-tracking related errors, such as halving- and doubling errors. In the case of 296 of the target nouns, mean pitch and pitch range were manually measured in Praat ([Boersma and Weenink, 2022](#)) and corrected in the data files.

5 Statistical analysis and results

We assessed the data using linear mixed-effects modeling in R Studio ([RStudio Team, 2020](#)) with the lme4 package ([Bates et al., 2015](#)). The outcome variables were duration, mean intensity,³ mean pitch and pitch range. The random factors were *Participant* and *Item*. Two sets of analyses were conducted for each outcome variable in the data in order to jointly address our research questions.

In the first set of analyses, we compared the nouns in the NSF and NOF conditions (where the subject noun was focused in the former but not in the latter, and vice versa for the object noun) to study the use of prosody in distinguishing narrow focus from non-focus, i.e., the effect of focus. The fixed factors included in the analyses were *Focus* (i.e., narrow focus or non-focus), *Position* (i.e., initial or final). In the second set of analyses, we examined the potential differences in the expression of contrastive focus and narrow focus, i.e., the effect of focus type. Since all target items were produced in object (final) position, the factor *Position* was no longer relevant. The trials used in this set of analyses were those in the NOF and NOCF conditions. The analyses only included *FocusType* (i.e., narrow focus or contrastive) as a fixed factor.

In both sets of analyses, to examine whether a specific prosodic cue was used to distinguish between two focus conditions, we built models using the aforementioned factors, in a stepwise fashion. Starting with a base model that only included the random factors (with no random slopes due to singular-fit issues), we then added the main effects of the fixed factor(s), and the two-way interaction between the two fixed factors (only in the first set of analyses). Model comparisons were conducted using the ANOVA function in RStudio, in order to derive the model with the best fit. We note that, for all results, we will only focus on statistically significant results concerning

participants produced in the answer. The question asked by the experimenter was of the form 'Est-ce que les jumelles jouent avec une pierre?' 'Are the twins playing with a stone?', which included the preposition 'avec/with'. However, most of the answers given were produced with another preposition 'au/at', even though the robot answer was of the form 'Les jumelles jouent avec un ballon' 'The twins play with a ball.'

³ Because the distance between the mouth of the participants and the microphone was not strictly controlled for adults and children, the mean intensity results reported here should be interpreted with caution.

² The problematic NOCF item was removed from analysis due to a mismatch between the preposition used in the question and the preposition that

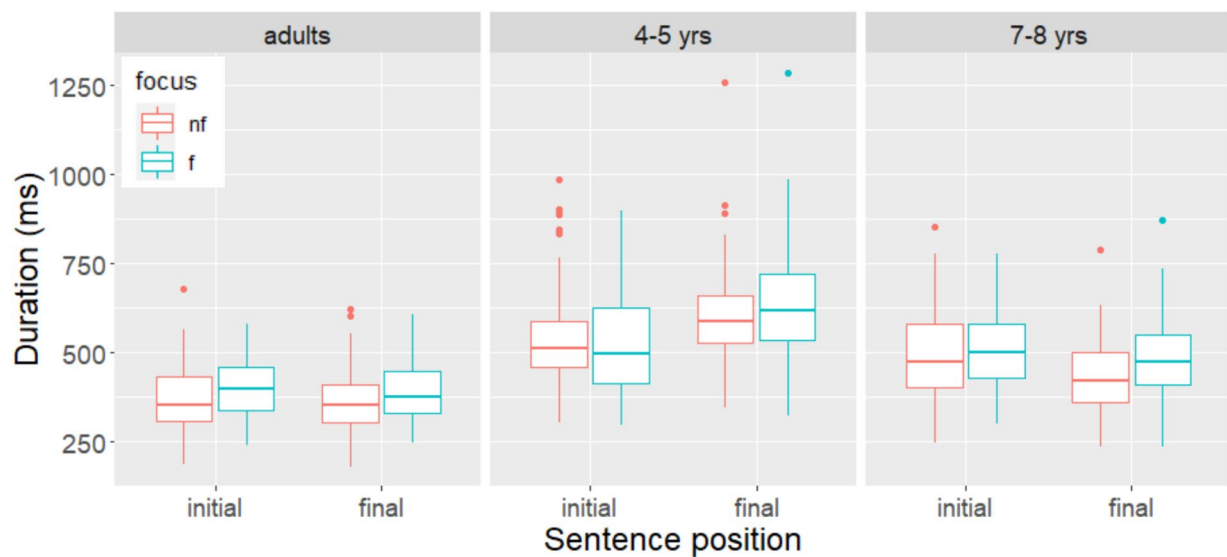


FIGURE 2

Duration on target nouns in non-focus and narrow focus, by group and sentence position. Non-focus is represented as "nf" (red); narrow focus as "f" (green).

the fixed factors *Focus* and *Focus type* and their interaction with another fixed factor.

5.1 The effect of focus: narrow focus vs. non-focus⁴

5.1.1 Adults

For **duration**, the model with the best fit showed only a significant main effect of *Focus* ($\beta = 30.322$, $SE = 6.072$, $t = 4.993$, $p < 0.001$). As shown in Figure 2, the nouns were produced with a longer duration in the narrow focus condition than in the non-focus condition, regardless of sentence position.

For **mean intensity**, the model with the best fit was the one that included both fixed factors and their interaction, revealing a significant interaction between *Focus* and *Position* ($\beta = 1.407$, $SE = 0.485$, $t = 2.89$, $p < 0.004$). Subsequent analysis showed that the main effect of *Focus* was only significant in final position ($\beta = 1.95$, $SE = 0.359$, $t = 5.42$, $p < 0.001$). Overall, the nouns were produced with a greater mean intensity in sentence-initial position than in sentence-final position; they were produced with a significantly higher mean intensity in the focused condition than in the non-focus condition only in sentence-final position, as shown in Figure 3.

Regarding **mean pitch**, the model with the best fit also showed a significant interaction between the fixed factors *Focus* and

Position ($\beta = 12.462$, $SE = 4.983$, $t = 2.501$, $p < 0.05$). Subsequent analysis revealed that the main effect of *Focus* was only significant in sentence-final position ($\beta = -15.772$, $SE = 3.748$, $t = -4.208$, $p < 0.001$). That is, the nouns were produced with a significantly higher mean pitch in the focus condition than in the non-focus condition only sentence-finally, as can be seen in Figure 4.

Finally, the model with the best fit only showed a significant main effect of *Position* for **pitch range** ($\beta = 10.166$, $SE = 4.011$, $t = 2.535$, $p < 0.05$).

To sum up, the adults made use of prosodic cues such as duration, mean intensity and mean pitch to signal narrow focus in sentence-final position, and only duration in sentence-initial position.

5.1.2 4- to 5-year-old French-speaking children

The model with the best fit retained a significant main effect of *Position* for the outcome variables **duration** ($\beta = 89.47$, $SE = 33.64$, $t = 2.65$, $p < 0.02$), **mean intensity** ($\beta = -2.341$, $SE = 0.55$, $t = -4.25$, $p < 0.001$), and **mean pitch** ($\beta = 26.042$, $SE = 2.158$, $t = 12.069$, $p < 0.001$). Thus, the nouns in sentence-initial position were produced differently from those in sentence-final position. Regardless of whether they were focused or not, they had a shorter duration, a higher intensity, and a higher mean pitch in sentence-initial position (Figures 2–4). The best-fit model for the outcome variable **pitch range** contained only random factors. There was no effects of *focus* or *position*.

5.1.3 7- to 8-year-old French-speaking children

The model with the best fit for the outcome variables **duration** retained a significant interaction between the fixed factors *Focus* and *Position* ($\beta = 42.22$, $SE = 16.29$, $t = 2.59$, $p < 0.001$). Subsequent analysis showed that the main effect of *Focus* was only statistically significant in sentence-final position ($\beta = 54.08$, $SE = 10.25$, $t = -5.274$, $p < 0.001$). Thus, the nouns were produced with a longer duration in the narrow focus condition than in the non-focus condition only sentence-finally, as shown in Figure 2.

⁴ Our subsequent exploratory mixed-effects logistic regression modeling with *Group* (4–5 years, 7–8 years, or adults) has revealed a significant interaction of *Group* \times *Focus* \times *Position* in the model with the best fit for duration ($p < 0.05$), mean intensity ($p < 0.05$), and mean pitch ($p < 0.05$), but not for pitch range. These results suggest that differences observed between age groups in the use of pitch range need to be taken with caution.

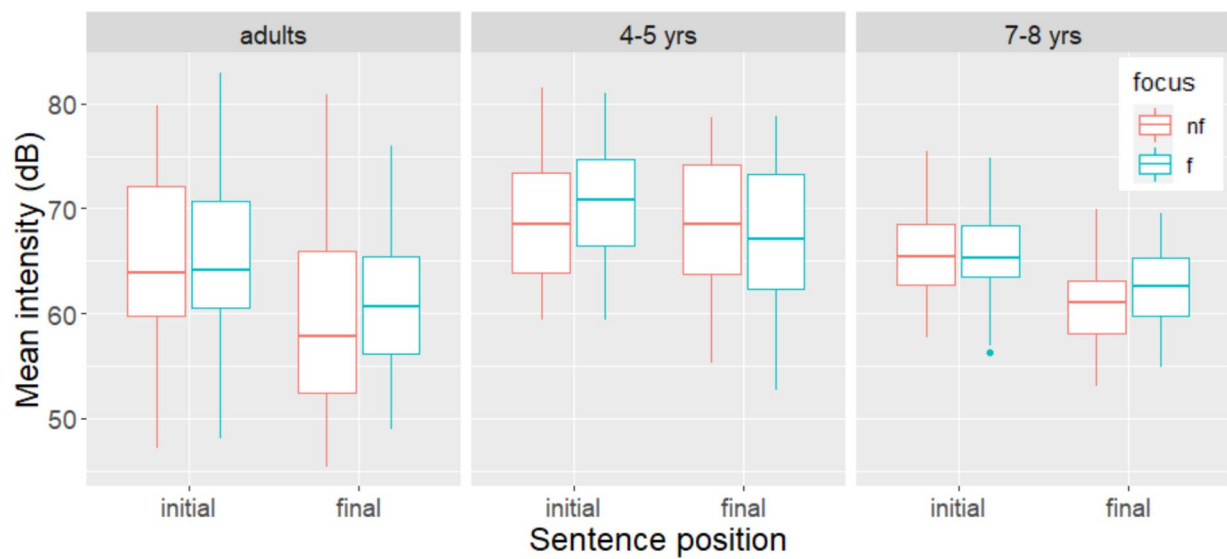


FIGURE 3

Mean intensity on target nouns in non-focus and narrow focus, by group and sentence position. Non-focus is represented as "nf" (red); narrow focus as "f" (green).

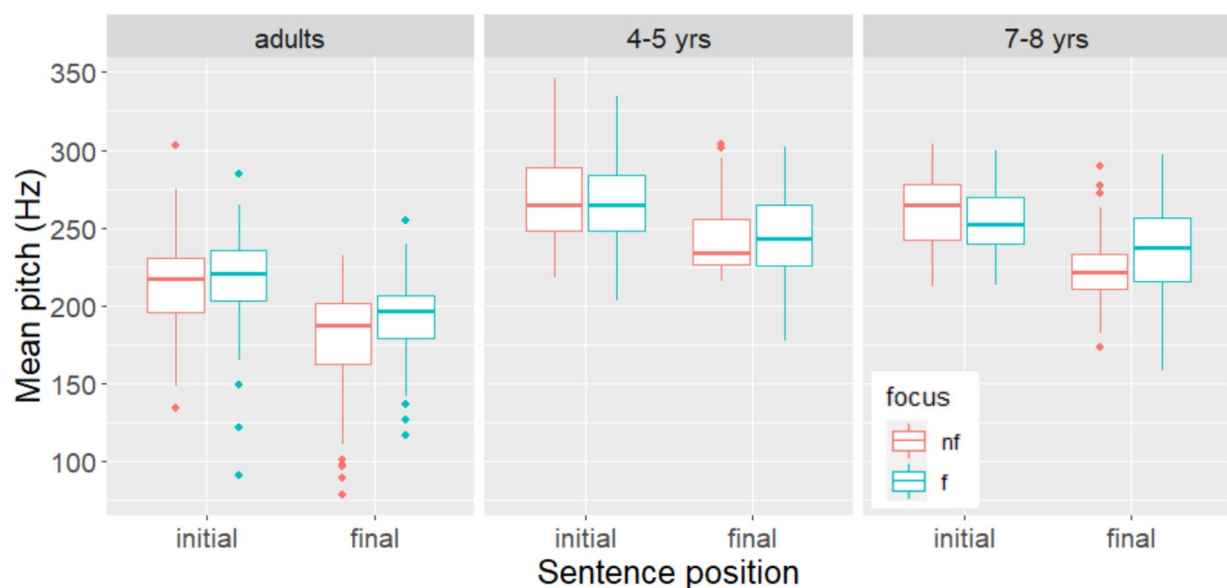


FIGURE 4

Mean pitch on target nouns in in non-focus and narrow focus, by group and sentence position. Non-focus is represented as "nf" (red); narrow focus as "f" (green).

The model with the best fit for the outcome variable **mean intensity** included a significant interaction between the fixed factors *Focus* and *Position* ($\beta = 1.816$, $SE = 0.54$, $t = 3.32$, $p < 0.001$). Subsequent analysis showed that the main effect of *Focus* was only significant in sentence-final position ($\beta = 1.85$, $SE = 0.4125$, $t = 4.5$, $p < 0.001$). Thus, the nouns were produced with similar intensity in the focus and non-focus conditions sentence-initially, but the nouns were produced with higher intensity in the focus condition than in the non-focus condition sentence-finally, as shown in Figure 3.

The model with the best fit for the outcome variable **mean pitch** only retained a marginally significant main effect of focus ($\beta = -3.467$, $SE = 1.822$, $t = -1.903$, $p = 0.058$). The nouns tended to be produced with a higher mean pitch when focused than when not focused, regardless of sentence position, as shown in Figure 4.

The model with the best fit for the outcome variable **pitch range** only showed a main effect of position ($\beta = -19.373$, $SE = 4.563$, $t = -4.245$, $p < 0.001$).

Taken together, these results show the 7- to 8-year-olds produced the sentence-final nouns with a longer duration and a higher intensity

in the focus condition than in the non-focus condition in sentence-final position. Further, they tended to produce the nouns with a higher mean pitch in the focus condition than in the non-focus condition regardless of sentence position.

5.2 The effect of focus type: contrastive focus vs. narrow focus⁵

For adults, the model with the best fit for the outcome variable **duration** showed a main effect of *FocusType* ($\beta = 23.06$, $SE = 7.54$, $t_{3,058}$, $p < 0.005$), such that the object nouns were produced with a shorter duration in contrastive focus than in narrow focus. No main effect of *FocusType* was found for the other outcome variables in the adult data.

No significant result appeared in the case of the 4- to 5-year-olds and the 7- to 8-year-olds, revealing that these participants did not use the prosodic cues investigated to distinguish between contrastive and narrow focus.

In sum, the findings indicate that the adults implemented only a decrease in duration to distinguish contrastive focus from narrow focus. The 4- to 5-year-olds and 7- to 8-year-olds did not differentiate these two focus types through prosodic cues.

6 Discussion and potential limitations

The findings on the French-speaking adults showed that duration, mean intensity and mean pitch were used to distinguish narrow focus from non-focus, similar to findings reported in previous studies on French. However, the results indicate an asymmetry in its realization between sentence-initial and final positions. While three prosodic parameters were used in sentence-final position, only duration was used to mark focus in sentence-initial position. This finding is in line with the observation that there are other non-prosodic strategies preferred by French-speakers for marking focus in sentence-initial position. Furthermore, the adults in our study produced the target words (i.e., the object nouns) with a shorter duration in the contrastive focus condition than in the narrow focus condition. This use of duration has not been reported before for French. It may be related to pragmatic reasons such as the intention to minimise potential negative connotation of correcting someone unfamiliar but of a similar age to oneself in the current set-up.

For the 4- to 5-year-olds, the current results have established that at this age, children do not implement focus through prosodic cues, in line with Hypothesis I. However, they were successful in manipulating these cues to distinguish sentence positions: similar

to what is observed in adults from prior literature, 4- to 5-year-olds produced sentence-initial nouns with a higher mean intensity and a higher mean pitch than sentence-final nouns. Contrasting from adults, they additionally implemented the parameter of duration but were unsuccessful in using pitch range for this purpose. This suggests that although 4- to 5-year-olds are unable to signal focus through prosodic cues, they may already use some of these cues for other purposes. However, as the sentence-initial nouns were segmentally different from the sentence-final nouns, a direct comparison in their duration, mean intensity and mean pitch is not possible. Future research compared the same nouns in different sentence positions is needed to validate this speculation.

In the 7- to 8-year-olds, the findings indicate that, at this stage, children have developed some ability to use prosody for focus marking purposes, supporting Hypothesis I. Although they did not use prosody to distinguish focus types, i.e., contrastive focus and narrow focus, they used an increase in duration and mean intensity and to a lesser degree, an increase in mean pitch to differentiate narrow focus from non-focus in sentence-final position, similar to the French-speaking adults in our study. Furthermore, the 7- to 8-year-olds appeared to use an increase in mean pitch for this purpose in sentence-initial position, unlike the French-speaking adults who used duration for this purpose. These results indicate faster acquisition of prosodic focus marking in subjects than in objects, supporting Hypothesis II. As has become clear from section 2, subject-focus is preferably marked via cleft structures in French. There is evidence that children acquire such structures in focus contexts very early on, i.e., around the age of 2 years (Labelle, 1990; Belletti, 2005; De Cat, 2007; Soares-Jesel and Lobo, 2019; Lahousse and Jourdain 2024). Thus, the lower incidence of canonical Subject-Verb-Object sentences in subject-focus contexts could be a factor contributing to the later acquisition of prosodic focus marking in subjects.

We conclude our discussion by raising potential limitations of this study and proposing opening avenues for future research. One potential limitation relates to the observed difficulty some 4- to 5-year-old children had with certain words. More specifically, these children audibly struggled with pronouncing “docteur” (*doctor*) or “bâtir” (*built*). In cases where the error was on the noun, the trial was omitted from the analyses. However, since the verb was not included in the analyses, errors on these trials were maintained (given that there was no additional issue). Although this may seem unproblematic for the relatively few trials concerned, it is possible that if several children audibly struggled with these items, other children may have also had difficulties. A second point of attention is related to the fact that we have examined prosodic focus marking in children speaking the standard variety of French spoken in France. Due to prosodic differences between regional varieties of French (Obin et al., 2012), it can be insightful to replicate our study in children speaking prosodically distinct regional varieties of French. Finally, it remains to be investigated at what age exactly children acquire fully adult-like abilities in prosodic focus marking. Future research is needed to examine prosodic focus marking in older children, such as 10- to 11-year-olds based on previous studies on children acquiring other languages.

⁵ Our subsequent exploratory mixed-effects logistic regression modeling with *Group* (4–5 years, 7–8 years, or adults) has revealed a significant interaction of *Group* \times *FocusType* in the model with the best fit for mean pitch ($p < 0.001$), but not for duration, mean intensity and pitch range. These results suggest that differences observed between age groups in the use of duration, mean intensity and pitch range need to be taken with caution.

7 Conclusion

This study makes a new empirical contribution by providing insights into the development in the use of prosody for marking focus in French-speaking children. Our findings corroborate past literature on French-speaking adults on their use of prosody in focus marking in the absence of syntactic means and on French-speaking 4- to 5-year-olds on their lack of use of prosody in focus marking, showing differences between French-speaking children and children acquiring a language relying on prosody or both prosody and syntax to a similar degree for focus marking. Further, it provides first evidence that 7- to 8-year-olds use prosody to mark focus in sentence-final position (objects) in a more adult-like manner than in sentence-initial position (subjects), arguably caused by a more dominant role of syntactic means in marking focus on sentence-initial subject in French. Together, our study sheds new light on the influence of relative importance of prosodic and non-prosodic means on the acquisition of prosodic focus marking. Specifically, it shows that syntax-dominance can not only influence the route of acquisition (Chen, 2018) but also the rate of acquisition of prosodic focus marking.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements at the time of testing. The participants or the legal guardians provided their written informed consent to participate in this study.

Author contributions

ED: Conceptualization, Funding acquisition, Investigation, Writing – original draft, Writing – review & editing. LL: Data curation, Formal analysis, Writing – original draft. AC: Conceptualization, Data curation, Formal analysis, Funding

References

- Arnhold, A., Chen, A., and Järviö, J. (2016). Acquiring complex focus-marking: Finnish four- to five-year-olds use prosody and word order in interaction. *Front. Psychol.* 7:1886. doi: 10.3389/fpsyg.2016.01886
- Bates, D., Machler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67:1–48.
- Belletti, A. (2005). Extended doubling and the VP periphery. *Probus*, 17. doi: 10.1515/prbs.2005.17.1.1
- Beysade, C., Hemforth, B., Marandin, J.-M., and Portes, C. (2009). “Prosodic marking of information focus in French” in *Actes d’interfaces, discours et prosodie*. eds. H.-Y. Yoo and E. Delais-Roussarie (Paris, France: Université Paris 7), 109–122.
- Beysade, C., Hemforth, B., Marandin, J.-M., and Portes, C. (2015). “Prosodic realizations of information focus in French” in *Explicit and implicit prosody in sentence processing*. eds. L. Frazier and T. Gibson (Heidelberg: Springer International Publishing), 39–61.
- Boersma, P., and Weenink, D. (2022). Praat: doing phonetics by computer [Computer program]. Version 6.2.06, retrieved 30 August 2022 from <http://www.praat.org/>
- Chen, A. (2011). Tuning information packaging: Intonational realization of topic and focus in child Dutch. *J. Child Lang.* 38, 1055–1083. doi: 10.1017/S0305000910000541
- Chen, A. (2018). Chapter 15. Get the focus right across languages. *Trends Lang. Acquis. Res.* 23:295–314. doi: 10.1075/tilar.23.15che
- Chen, A., and DestrUEL, E. (2010). Intonational encoding of focus in Toulousian French. *Proceedings of Speech Prosody 2010*. doi: 10.21437/SpeechProsody.2010-165
- Chen, A., Esteve-Gibert, N., Prieto, P., and Redford, M. (2020). “Development in phrase-level from infancy to late childhood” in *The Oxford handbook of language prosody*. eds. C. Gussenhoven and A. Chen (Oxford: Oxford University Press).

acquisition, Methodology, Project administration, Supervision, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the research funding of the Max Planck Institute for Psycholinguistics and a VICI grant awarded to AC by the Dutch Research Council (NWO) (grant number VI.C.201.109).

Acknowledgments

We would like thank participating children from the Puygouzon Primary School in Albi, and adults in Toulouse, France for their help in this study, Christine Gipper for assistance in data collection, Helen Sprengers for assistance in making Figure 1, and Luna Scheijmans for assistance in acoustic analysis. Preliminary results of this study were reported in LL’s unpublished MA thesis completed under the supervision of AC at Utrecht University in 2023.

Conflict of interest

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.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2024.1360308/full#supplementary-material>

- Chen, A., and Höhler, B. (2018). Four- to five-year-olds use of word order and prosody in focus marking in Dutch. *Linguist. Vang.* 4:20160101. doi: 10.1515/lingvan-2016-0101
- Clech-Darbon, A., Rebuschi, G., and Rialland, A. (1999). "Are there cleft sentences in French" in *The grammar of focus*. eds. L. Tuller and G. Rebuschi (Amsterdam: Benjamins), 83–118.
- De Cat, C. (2007). *French dislocation: Interpretation, syntax and acquisition*. Oxford, UK: Oxford University Press.
- D'Imperio, M., German, J. S., and Michelas, A. (2012). "A multi-level approach to focus, phrasing and intonation in French" in *Prosody and meaning*. eds. G. Elordieta and P. Prieto (Berlin: Mouton de Gruyter), 11–34.
- Destruel, E. (2013). An empirical approach to the pragmatics of the (non)-canonical French sentences. Proceedings from the 46th Chicago linguistic society conference, Chicago, USA, pp. 43–58.
- Di Cristo, A. (1996). Intonation in French. In D. Hirst and Cristo A. Di (Eds.), *Intonation systems: A survey of twenty languages* (pp. 195–218). Cambridge: Cambridge University Press.
- Di Cristo, A., and Hirst, D. (1993). Rythme syllabique, rythme mélodique et représentation hiérarchique de la prosodie du français. *Travaux de l'Institut de Phonétique d'Aix-en-Provence* 15, 9–24.
- Dohen, M., and Loevenbruck, H. (2004). Pre-focal rephrasing, focal enhancement and post-focal deaccentuation in French. Proceedings of the 8th international conference on spoken language processing, pp. 785–788.
- Esteve-Gibert, N., Loevenbruck, H., Dohen, M., and D'Imperio, M. (2021). Preschoolers use head gestures rather than prosodic cues to highlight important information in speech. *Dev. Sci.* 25:e13154. doi: 10.1111/desc.13154
- Féry, C. (2001). "Focus and phrasing in French" in *A Festschrift for Arnim von Stechow*. eds. C. Féry and W. Sternefeld (Berlin. Française: Akademie Verlag), 153–181.
- German, J. S., and D'Imperio, M. (2016). The status of the initial rise as a marker of focus in French. *Lang. Speech* 59, 165–195. doi: 10.1177/0023830915583082
- Gussenhoven, C. (2007). "Types of focus in English" in *Topic and focus: Cross-linguistic perspectives on meaning and intonation*. eds. C. M. Lee, M. Gordon and D. Büring (Dordrecht: Springer Publishing), 83–100.
- Hornby, P. A., and Hass, W. A. (1970). Use of contrastive stress by preschool children. *J. Speech Lang. Hear. Res.* 13, 395–399. doi: 10.1044/jshr.1302.395
- Ito, K. (2018). "Gradual development of focus prosody and affect prosody comprehension: a proposal from a holistic approach" in *Prosodic development in first language acquisition*. eds. P. Prieto and N. Esteve-Gibert (Amsterdam: John Benjamins), 295–314.
- Jun, S.-A., and Fougeron, C. (2000). "A phonological model of French intonation" in *Intonation: Analysis, modelling and technology*. ed. A. Botinis (Boston: Kluwer Academic), 209–242.
- Kiss, K. E. (1998). Identificational focus versus information focus. *Language* 74, 245–273. doi: 10.1353/lan.1998.0211
- Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica* 55, 243–276. doi: 10.1556/ALing.55.2008.3-4.2
- Kügler, F., and Calhoun, S. (2020). Prosodic encoding of information structure. In C. Gussenhoven and A. Chen (eds.), *The Oxford Handbook of Language Prosody*, 453–467. Oxford University Press.
- Labelle, M. (1990). Predication, wh-movement, and the development of relative clauses. *Language Acquisition*, 1, 95–119.
- Ladd, D. R., and Arvaniti, A. (2023). Prosodic Prominence Across Languages. *Annual Review of Linguistics*, 9, 171–193. doi: 10.1146/annurev-linguistics-031120
- Lahousse, K., and Jourdain, M. (2024). The emergence and early development of c'est 'it is' clefts in French L1. In: *It-Clefts. Empirical and Theoretical Surveys and Advances*. C. Bonan and A. Ledgeway (Eds.), Berlin, Boston: De Gruyter Mouton, pp. 135–156
- Lambrech, K. (1994). *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge, UK: Cambridge University Press.
- Lee, Y. C. (2015). Prosodic focus within and across languages. Publicly Accessible Penn Dissertations. 1534. Available at: <https://repository.upenn.edu/edissertations/1534>
- Machač, P., and Skarnitzl, R. (2009). Phonetic analysis of parasitic speech sounds. In: Proceedings of the 19th Czech-German Workshop—Speech Processing, Prague.
- Ménard, L., Loevenbruck, H., and Savariaux, C. (2006). Articulatory and acoustic correlates of contrastive focus in French children and adults. In J. Harrington and M. Tabain. *Speech Production: Models, Phonetic Processes and Techniques* (pp. 227–251). New York, USA: Psychology Press.
- Michelas, A., Faget, C., Portes, C., Lienhart, A.-S., Boyer, L., Lançon, C., et al. (2014). Do patients with schizophrenia use prosody to encode contrastive discourse status? *Front. Psychol.* 5:755. doi: 10.3389/fpsyg.2014.00755
- Obin, N., Steien, G. B., Avanzi, M., Bardiaux, A., and Bordal, G. (2012). Speech prosody of French regional varieties. Proceedings of Speech Prosody 2012. ISCA Archive. https://www.isca-archive.org/speechprosody_2012/avanzi12b_speechprosody.pdf
- Portes, C., and Reyle, U. (2022). Combining syntax and prosody to signal information structure: the case of French. Proceedings of the International Conference on Speech Prosody, 2022-May, 87–91.
- Romøren, A. S. H. (2016). *Hunting highs and lows: The acquisition of prosodic focus marking in Swedish and Dutch*. Dissertation. The Netherlands: LOT publications.
- Romøren, A. S. H., and Chen, A. (2022). The acquisition of prosodic marking of narrow focus in central Swedish. *J. Child Lang.* 49, 213–238. doi: 10.1017/S0305000920000847
- Rooth, M. (1992). A theory of focus interpretation. *Nat. Lang. Semant.* 1, 75–116. doi: 10.1007/BF02342617
- Rossi, M. (1985). L'intonation et l'organisation de l'énoncé. *Phonetica* 42, 135–153. doi: 10.1159/000261744
- RStudio Team. (2020). *RStudio: Integrated development for R*. RStudio, PBC, Boston, MA. <http://www.rstudio.com/>
- Sauermann, A., Höhle, B., Chen, A., and Järvi, J. (2011). "Intonational marking of focus in different word orders in German children" in *Proceedings of the 28th west coast conference on formal linguistics*. eds. M. B. Washburn, K. Mc Kinney-Bock, E. Varis, A. Sawyer and B. Tomaszewicz (Somerville, MA: Cascadia Proceedings Project), 313–322.
- Schwarz, A. (2009). Tonal focus reflections in Buli and some Gur relatives. *Lingua*, 119, 950–972.
- Soares-Jesel, C., Lobo, M. (2019). The acquisition of clefts: syntax and information structure. (Paper presented at the International Workshop on the L1 and L2 acquisition of information structure, Katholieke Universiteit Leuven, 25–26 April 2019.)
- Szendrői, K., Bernard, C., Berger, F., Gervain, J., and Höhle, B. (2017). Acquisition of prosodic focus marking by English, French, and German three-, four-, five- and six-year-olds. *J. Child Lang.* 45, 219–241. doi: 10.1017/S0305000917000071
- Touati, P. (1987). *Structures prosodiques du suédois et du français*, Working Paper 21.; Lund, Germany: Lund University Press.
- Vallduv, E., and Engdahl, E. (1996). The linguistics realization of information packaging. *Linguistics*, 3, 459–520.
- Wonnacott, E., and Watson, D. G. (2008). Acoustic emphasis in four year olds. *Cognition* 107, 1093–1101. doi: 10.1016/j.cognition.2007.10.005
- Xu, Y. (2013). ProsodyPro – A tool for large-scale systematic prosody analysis. In Bigi, B., Hirst, D., Lavaud, J., and Pichon-Starke, C. (Eds.) Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013) (pp. 7–10). Aix-en-Provence, France: Laboratoire Parole et Langage (LPL).
- Yang, A., and Chen, A. (2018). The developmental path to adult-like prosodic focus-marking in mandarin Chinese-speaking children. *First Lang.* 38, 26–46. doi: 10.1177/0142723717733920
- Zimmermann, M., and Onea, E. (2011). Focus marking and focus interpretation. *Lingua* 121, 1651–1670. doi: 10.1016/j.lingua.2011.06.002



OPEN ACCESS

EDITED BY

Joseph V. Casillas,
Rutgers, The State University of New Jersey,
United States

REVIEWED BY

Yong-cheol Lee,
Hannam University, Republic of Korea
Hyunah Baek,
Ajou University, Republic of Korea
Hanyong Park,
University of Wisconsin–Milwaukee,
United States

*CORRESPONDENCE

Anqi Yang
✉ anqi.yang@tju.edu.cn
Aoju Chen
✉ aoju.chen@uu.nl

RECEIVED 07 December 2023

ACCEPTED 30 May 2024

PUBLISHED 31 July 2024

CITATION

Yang A, Cho T, Kim S and Chen A (2024)
Prosodic focus marking in Seoul
Korean-speaking children: the use of
prosodic phrasing.
Front. Psychol. 15:1352280.
doi: 10.3389/fpsyg.2024.1352280

COPYRIGHT

© 2024 Yang, Cho, Kim and Chen. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Prosodic focus marking in Seoul Korean-speaking children: the use of prosodic phrasing

Anqi Yang^{1*}, Taehong Cho², Sahyang Kim³ and Aoju Chen^{4*}

¹School of Foreign Languages, Tianjin University, Tianjin, China, ²Hanyang Institute for Phonetics and Cognitive Sciences of Language, Hanyang University, Seoul, Republic of Korea, ³Department of English Education, Hongik University, Seoul, Republic of Korea, ⁴Institute for Language Sciences, Utrecht University, Utrecht, Netherlands

Introduction: Prosodic focus marking in Seoul Korean is known to be achieved primarily through prosodic phrasing, different from the use of prosody for this purpose in many other languages. This study investigates how children use prosodic phrasing for focus-marking purposes in Seoul Korean, compared to adults.

Methods: Using a picture-matching game, we elicited semi-spontaneous production of SOV sentences in various focus conditions from monolingual Seoul Korean-speaking children aged 4 to 11 years.

Results: We found that the children varied prosodic boundaries to distinguish narrow focus from pre-focus and broad focus in a largely adult-like manner at the age of 4 to 5; at this age, they did not distinguish narrow focus from post-focus or contrastive focus using prosodic boundaries, similar to the adults. Their use of the prosodic boundaries in distinguishing the focus conditions was not fully adult-like in terms of frequency until the age of 10 to 11.

Discussion: In conjunction with the findings of previous studies on the acquisition of focus marking in Germanic languages, performed using a similar experimental method, our findings suggest that Seoul Korean-speaking children acquire the use of prosodic phrasing earlier than Dutch-speaking children acquiring the use of pitch accent but slightly later than Stockholm Swedish-speaking children acquiring the use of a prominence-marking high tone. These findings imply that the rate of focus-marking acquisition depends on the transparency of the form-meaning mapping between the phonological cue and focus.

KEYWORDS

acquisition, focus, prosodic phrasing, prosodic marking of focus, Seoul Korean

1 Introduction

Previous studies on the acquisition of prosodic focus marking in various languages have revealed that children acquiring a language that relies on prosody or both prosody and syntax for focus marking show adult-like use of at least some language-specific prosodic focus-marking cues by the age of 5 and further develop this ability until the age of 10 or 11 (e.g., Hornby and Hass, 1970; MacWhinney and Bates, 1978; Wells et al., 2004; Thorson and Morgan, 2020, on English; Chen, 2009, 2011a,b; Romøren, 2016, on Dutch; Grünloh et al., 2015, on German; Romøren and Chen, 2021, on Stockholm Swedish, hereafter Swedish; Yang and Chen, 2018, on Mandarin). However, the specific developmental trajectory differs for children acquiring different languages due to typological differences in prosodic system and prosodic focus marking (Chen, 2018). The present study is concerned with children learning Seoul Korean (hereafter Korean), a language that differs from more widely studied languages

such as Dutch, English, Finnish, German, Swedish, and Mandarin in that it primarily uses prosodic phrasing in focus marking (Jun, 1993, 2007b; Jun and Lee, 1998).

Focus is prosodically encoded in many languages (e.g., Vallduví and Engdahl, 1996; Gussenhoven, 2007; Kügler and Calhoun, 2021). Among the most frequently studied focus types in literature, narrow focus (i.e., focus on a word of a syntactic constituent, like “the bread” in (1B)) differs from narrow contrastive focus (i.e., narrow focus carrying contrast¹, hereafter contrastive focus, like “the bread” in (2B)) in contrastivity, and differs from broad focus (i.e., focus over a constituent larger than a word, like (3B)) in size of focal constituent. Generally, a word is realized with more prosodic prominence in narrow focus and contrastive focus than when not focused or in broad focus. Contrastive focus can be realized with additional prosodic prominence, compared to non-contrastive narrow focus, in certain speech styles in some languages (e.g., read speech in Mandarin) (Chen and Braun, 2006).

- (1) A: Look! The dog², and it holds a painting brush. It looks like the dog draws something. What does the dog draw?

B: 개가	[빵을] ³	그려요.
keka	p*anjil	kiljɔjo.
The dog	[the bread]	draws.

- (2) A: Look! The bear. The bear looks a bit puzzled. It looks like the bear looks for something. I will make a guess: The bear looks for the egg.

B: 곰이	[빵을]	찾아요.
komi	p*anjil	tʃʰatʃajo.
The bear	[the bread]	looks for.

- (3) A: Look! My picture is very blurry. I cannot see anything clearly. What happens in your picture?

B: [말이	빵을	그려요].
mali	p*anjil	kiljɔjo.
[The horse	the bread	draws].

Regarding the precise prosodic means for achieving prominence, we distinguish *phonetic implementation* (hereafter *phonetic means*) and *phonological means*, following Chen (2009, 2011b, 2018). Specifically, some languages (e.g., Mandarin and Cantonese) rely on phonetic means. Speakers of these languages vary the phonetic implementation

of phonological categories such as lexical tones in the dimensions of duration, pitch, and intensity for focus-marking purposes, without changing the tonal identity of relevant words (e.g., Xu, 1999, on Mandarin; Wu and Xu, 2010, on Cantonese). For example, in Mandarin, a word in narrow focus is produced with a longer duration, wider pitch span, and higher intensity than the same word in non-focus, while its tonal category remains intact (e.g., Shih, 1988; Xu, 1999). Other languages (e.g., English, German, Dutch, Swedish, and Korean) primarily use *phonological means* to realize prosodic prominence. That is, speakers make coarse-grained changes in duration, pitch, and intensity that lead to a change in the phonological category of prosody. For example, in English, German, and Dutch, speakers can either accent words with certain types of pitch accents (e.g., rising vs. falling) or not accent words to distinguish narrow focus from non-focus (e.g., Gussenhoven, 2004, 2007; Baumann et al., 2006; Hanssen et al., 2008; Chen, 2009, 2011b); in Swedish, speakers can either assign or not assign a prominence-marking high tone to the end of a word for this purpose (Bruce, 2007; Romøren and Chen, 2021). However, these languages can differ in the transparency of the form-meaning mapping between the phonological cue and focus, i.e., how consistent the mapping is. For example, in Swedish, only focused words are produced with a word-final high tone. The mapping between the placement of the prominence-marking high tone and focus is thus highly transparent. By contrast, in Dutch, both focused and non-focused words can be accented with the same type of pitch accent, e.g., a falling accent in sentence-initial subject-noun phrases and a downstepped falling accent in sentence-final object-noun phrases, regardless of focus status. There is thus no consistent or transparent mapping between accentuation and focus (Chen, 2018). In such cases, speakers vary the phonetic implementation of pitch accents to distinguish focus and non-focus (Chen, 2009).

Differences in the transparency of the form-meaning mapping between the phonological cue and focus can lead to differences in the rate of acquisition in prosodic focus marking across languages (Chen, 2018). Previous studies on children acquiring a West Germanic language, which is relatively less transparent as discussed above, have shown that while children can already use accentuation to mark focus by the age of 5, their choice of accent type is not fully adult-like until the age of 7 or 8 (e.g., Hornby and Hass, 1970; MacWhinney and Bates, 1978; Wells et al., 2004; Thorson and Morgan, 2020, on English; Grünloh et al., 2015, on German; Chen, 2011a,b, on Dutch). In contrast, in Swedish, which is more transparent in the form-meaning mapping, phonological focus-marking is acquired earlier (Romøren, 2016; Romøren and Chen, 2021). That is, Swedish-speaking children are by and large adult-like at the age of 4 or 5 in assigning a prominence-marking high tone to the end of the word in narrow focus and contrastive focus conditions. The earlier acquisition of phonological focus marking in Swedish than in the West Germanic languages has been attributed to the more transparent form-meaning mapping between the phonological cue and focus in Swedish (Chen, 2018).

In the present study, we aim to extend the current understanding of the effect of transparency of form-meaning mapping by investigating how Korean-speaking children acquire the use of the phonological cue for focus-marking purposes. Korean is different from the previously studied languages in both prosodic system and prosodic focus marking. Regarding prosodic system, Korean has no word-level use of prosody. It is often classified as an edge-prominence

1 Focus can carry contrastive information, e.g., a correction to a certain piece of information introduced previously or an alternative to what has been mentioned already. While the focal items are regarded as being in contrastive focus in both conditions, the former is also referred to as corrective focus (Gussenhoven, 2004, 2007). In the present study, we use “contrastive focus” instead of “corrective focus” to refer to the former condition.

2 The examples are from the experimental materials of the present study. The referents are referred to with the definite article in the English glossary because they had been introduced in a picture-naming task before the production experiment.

3 The focused constituents are in square brackets in the examples.

language (Jun, 2014). That is, pitch movement in this language is aligned to prosodic phrases (Jun, 1993, 1998, 2000, 2005, 2014). Accentual Phrase (AP) is the smallest unit carrying a phrasal tone sequence, THLH, with the initial tone (T) being realized as either a high tone (H) or a low tone (L) at the left edge, depending on the laryngeal feature of the initial segment of the AP. The second H tone is generally realized on the second syllable of an AP, but is sometimes realized on the third syllable when an AP is longer than four syllables; when an AP contains fewer than four syllables, one or both of the two middle tones may be undershot, with the choice of tones undershot varied across speakers, discourse contexts, and other linguistic factors (Jun, 2005). The final H tone is realized on the last syllable and marks the right edge of an AP. In addition to the tonal marking at both edges, AP is also marked by domain-initial (segmental) strengthening at the left edge (Cho, 2022; Kim et al., 2024). An AP consists of one or more Phonological Words (PWs). An Intonational Phrase (IP) consists of one or more APs and is marked by a phrase-final boundary tone and phrase-final lengthening (or pre-boundary lengthening) at the right edge (Cho, 2022; Kim et al., 2024).⁴

Regarding prosodic focus marking, past work based on read speech shows that Korean primarily uses prosodic phrasing for this purpose (e.g., Jun, 1993, 2007b). A word in a narrow focus or contrastive focus condition typically initiates a prosodic phrase, which can be either an AP or an IP, with the following words tending to be integrated into the same phrase as the focused word, resulting in dephrasing (Jun, 1993; Jun and Lee, 1998; Oh, 1999; Jun and Kim, 2007; Jeon and Nolan, 2017). Given that a phrasal boundary is either present or not (i.e., a discrete concept), the use of prosodic phrasing can be considered a phonological means of focus marking, like accent placement and choice of accent type in Dutch and the placement of a prominence-marking high tone in Swedish. However, prosodic phrasing, including dephrasing, can be influenced by factors other than focus marking, such as speech rate, rhythm, semantic weight, length and syntactic structure of the utterance (cf. Jun, 1993, 2011). For example, when a short syntactic phrase (e.g., a verb phrase or noun phrase) or a short sentence as a whole is focused, each word in it tends to form an AP (Jun et al., 2006; Kim et al., 2006; Jun and Kim, 2007; Jun, 2011). Dephrasing does not play a role in focus marking in this case. Hatcher et al. (2023) have also demonstrated instances where focus realization can occur phrase-medially in Korean, notably without resorting to phrasing. The transparency of the form-meaning mapping between phonological cues and focus conditions in Korean can thus be considered lower, compared to Swedish. It may not be different from that in a West Germanic language, like Dutch,⁵ because both phrasing (in Korean) and pitch accent (in a West Germanic language) can occur in non-focused contexts.

However, prosodic phrase boundaries, especially IP boundaries, appear to be relatively easy to perceive in continuous speech streams, compared to the perception of prosodic prominence associated with pitch accent. This has been shown to be the case of both linguistically trained adult listeners (e.g., Grice et al., 1996, on German; Jun et al., 2000, on Korean; Escudero et al., 2012, on Catalan) and naïve adult listeners with no prior linguistic knowledge presented with an unfamiliar language for the first time (Cole and Shattuck-Hufnagel, 2016; Cole et al., 2017; Bishop et al., 2020). Development literature suggests a very early ability to perceive and process prosodic phrase boundaries. For example, French-learning infants exhibit sensitivity to prosodic boundaries at birth (Christophe et al., 1994, 2001). Infants learning a West Germanic language change from relying on all possible cues in perception of major prosodic phrase boundaries to a subset of the cues between 4 and 8 months, partially reflecting the relative importance of the cues in the target language (Seidl, 2007; Johnson and Seidl, 2008; Seidl and Cristià, 2008). They exhibit adult-like processing of major prosodic phrase boundaries in the brain at 6 months (Holzgrefe-Lang et al., 2018). These findings suggest that a prosodic boundary may be a perceptually more recognizable cue to focus than accentuation, making the mapping between a prosodic phrase boundary and focus possibly more transparent in Korean than the mapping between accentuation and focus in a West Germanic language to young language learners. Notably, children appear to be similar in their development in the production of pitch accents in a West Germanic language and the production of AP pitch patterns and IP-final boundary tones in Korean, independent of the mapping with meaning like focus. For example, by the age of 2, children acquiring a West Germanic language can produce the core of an adult-like inventory of pitch accents (Chen et al., 2020); Korean-speaking children can produce certain AP pitch patterns and IP-final boundary tones (Jun, 2007a).

According to Chen's (2018) cross-linguistic model of acquisition of prosodic focus marking, higher transparency of the form-meaning mapping between focus and the phonological cue will lead to a faster rate of acquisition in prosodic focus marking. If this holds, we hypothesize that Korean-speaking children will exhibit adult-like phonological focus marking in some or all aspects relatively later than their Swedish-speaking peers, who are adult-like at the age of 4 to 5 years, but earlier than their Dutch-speaking peers, who are not fully adult-like yet at the age of 7 to 8 years.

To test this hypothesis, we have examined how Korean-speaking children aged 4 to 11 vary prosodic phrasing to distinguish (1) Narrow focus from non-focus (Effect of focus); (2) Narrow focus from broad focus (Effect of focal constituent size); and (3) Narrow focus from contrastive (narrow) focus (Effect of contrastivity), in comparison to Korean-speaking adults. Regarding the effect of focus, we predict that Korean-speaking children will be fully adult-like at the age of 7 to 8 years or later in more frequently using an AP and/or IP boundary before a word and a PW boundary after it when the word is in the narrow focus condition than in the non-focus conditions (i.e., pre-focus or post-focus). Regarding the effect of focal constituent size, we predict that the children will be fully adult-like at the age of 7 to 8 years or later in frequently using an AP and/or IP boundary before a word in both the narrow and broad focus conditions, and in more frequently using a PW boundary after a word when it is under narrow focus than broad focus. Regarding the effect of contrastivity, as previous studies on Korean reveal no evidence for the use of prosodic

⁴ An Intermediate Phrase (ip), which is larger than an AP but smaller than an IP, was added in a later version of the Korean prosodic hierarchy (Jun, 2007b, 2014). However, as ip is not clearly defined, transcribing ips in speech data is not straightforward, and its role in focus marking is not examined in the present study.

⁵ The present study was done using an adapted method from Chen (2011a,b), parallel to Romøren and Chen (2021). The results may thus be more comparable to Chen (2011a,b) on Dutch and Romøren and Chen (2021) on Swedish than to other previous studies.

phrasing to mark contrastivity and there is no language-independent reason to use prosodic phrasing to mark a contrast, we predict that Korean-speaking children will not vary prosodic boundaries to distinguish narrow focus from contrastive focus, like adults and this may be observable at the age of 4 to 5 years.

2 Method

2.1 Participants

Three groups of children participated in the experiment, including six 4- to 5-year-olds (average age: 5;3, range: 4;10–5;10), eight 7- to 8-year-olds (average age: 8;0, range: 7;4–8;10), and eight 10- to 11-year-olds (average age: 10;10, range: 10;3–11;11). They were recruited via Hanyang Institute for Phonetics and Cognitive Sciences of Language in Seoul, and came from diverse social-economic backgrounds. Twelve adults (six females and six males, average age: 24 years, range: 19–28 years) participated as a control group. They were students of Hanyang University at the time of testing. All participants spoke Seoul Korean as their native language and did not have any known language and/or cognitive impairment.

2.2 The picture-matching game

We adapted the picture-matching game used in [Chen \(2011b\)](#) to elicit semi-spontaneous production of sentences. The same game was used in recent studies on prosodic focus marking in children acquiring other languages ([Romøren, 2016](#); [Liu, 2017](#); [Yang and Chen, 2018](#); [Romøren and Chen, 2021](#)). In this game, the child was supposed to help the experimenter put pictures in matched pairs ([Figure 1](#)). Three piles of pictures were used. The experimenter and

the child each held a pile of pictures. The third pile lay on the table in a seemingly messy fashion. The experimenter's pictures always missed some information (e.g., the subject, the action, the object, or all three). The child's pictures always contained all three pieces of information. On each trial, the experimenter showed one of her pictures to the child, described the picture and asked a question about it or made a remark about the missing information (in the contrastive focus condition). The child then took a look at the corresponding picture in his/her pile and responded to the experimenter's question or remark. The experimenter then looked for the right picture in the messy pile and matched it with her own picture to form a pair.

As rules of the game, the child was asked to answer the experimenter's questions in full sentences and not to reveal his/her pictures to the experimenter. We constructed an experiment protocol outlining each step of the game, specifying the experimenter's instructions and responses for each trial. This protocol ensured consistent conduct of the experiment for all children and the provision of sufficient background information before each question or remark. The experimenter was instructed to follow the protocol closely but was encouraged to make spontaneous remarks that did not affect the information structure of the child's responses for the purpose of facilitating the interaction. Prior to the picture-matching game, a picture-naming task was conducted to ensure that the child would use the intended words to refer to the entities in the pictures. This procedure also rendered all the entities in the pictures referentially accessible.

2.3 Experimental materials

Sixty question-answer dialogues were embedded in the picture-matching game to elicit 60 SOV sentences with the sentence-medial target object-nouns in five focus conditions, as given below:

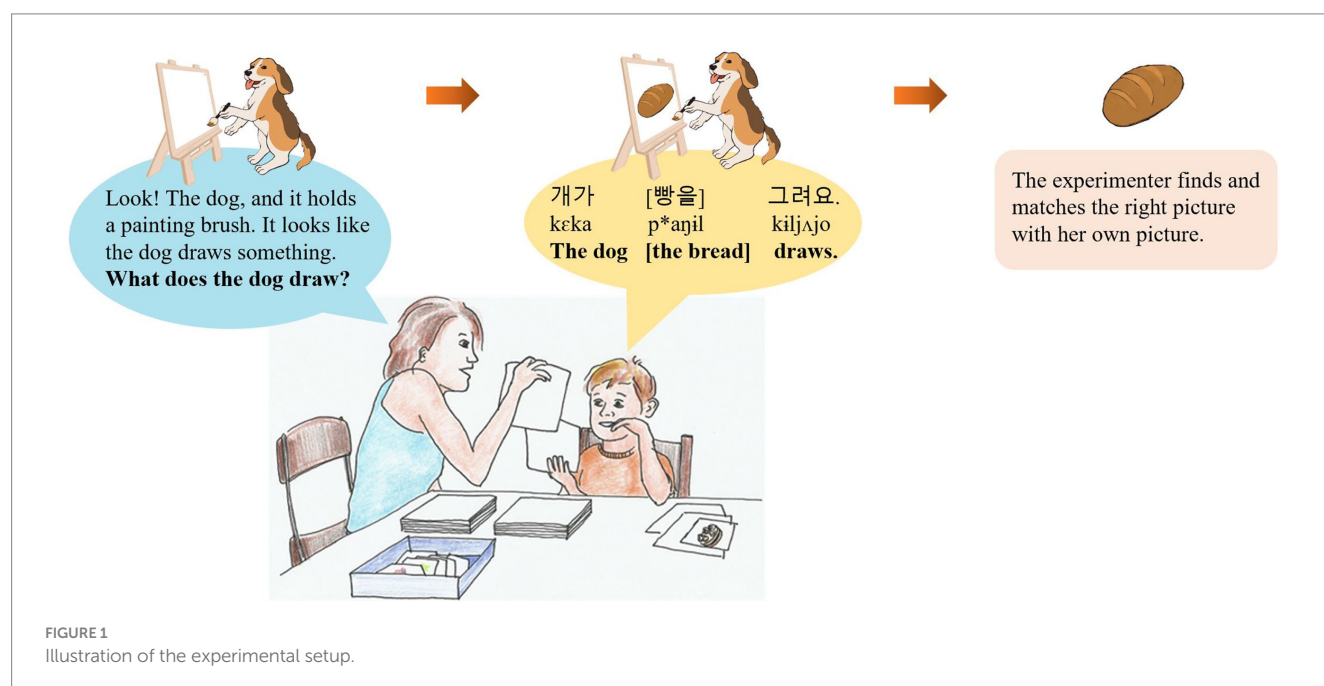


FIGURE 1
Illustration of the experimental setup.

- Narrow focus condition: when narrow focus was on the sentence-medial target object-noun, responding to a *what*-question, as illustrated in (1) in the introduction;
- Pre-focus condition: when narrow focus was on the sentence-final verb, responding to a *what-does-X-do-to-Y*-question, as in (4) below;
- Post-focus condition: when narrow focus was on the sentence-initial subject-noun, responding to a *who*-question, as in (5) below;
- Contrastive focus condition: when contrastive focus was on the sentence-medial target object-noun, correcting the experimenter's remark about the object, as in (2) in the introduction;
- Broad focus condition: when broad focus was over the whole sentence, responding to *what-happens* questions, as in (3) in the introduction.

(4) A: Look! The rat, and the bread. It looks like the rat does something to the bread. What does the rat do to the bread?

B: 쥐가	빵을	[만져요].
tɕwika	p*anil	mantɕajɔ.
The rat	the bread	[touches].

(5) A: Look! The bread, and someone looks a bit puzzled. It looks like someone looks for the bread. Who looks for the bread?

B: [소가]	빵을	찾아요.
soka	p*anil	tɕʰatɕajɔ.
[The cow]	the bread	looks for.

Note that narrow focus was included in three sentence-positions to make it possible to examine the effect of narrow focus on the sentence-medial object-nouns compared to the same words in the pre-focus and post-focus conditions, following previous studies on prosodic focus marking (e.g., [Chen, 2009](#)). Moreover, comparing the object-nouns in the narrow focus, contrastive focus and broad focus conditions allowed us to examine the prosodic difference between different focus types.

The 60 SOV sentences were unique combinations of five subject-nouns, 12 target object-nouns, and three verbs ([Table 1](#)). Each subject-noun was a monosyllabic lexical word followed by a nominative case marker, ⁶/ka/ or /i/. The target object-nouns included six two-syllable (or “short”) words and six four-syllable (or “long”) words, because APs with two to four syllables tend to occur frequently in Korean. The length of the target object-nouns was thus systematically varied to control for any potential effects. Each “short” word was a monosyllabic lexical word followed by an accusative case marker, /il/ or /lil/. As for the “long” words, except for /k*amakwilil/ (crows), which consisted of a three-syllable lexical word and the accusative case marker /lil/, each of the other words consisted of a disyllabic lexical word, a monosyllabic suffix /til/ indicating the plural form of the lexical word, and the accusative case marker /il/. Each target word was initiated with either a high-tone-triggering aspirated stop (i.e., /p^h/ or /k^h/) or fortis stop

(i.e., /p*/ or /k*/) or a low-tone-triggering lenis stop (i.e., /p/ or /k/) or a vowel (i.e., /a/), so that there would be varied AP tonal patterns in the data. Each target word appeared once in five focus conditions (12 target words × five focus conditions), leading to 60 “sentences” but without subject-nouns and verbs. The five subject-nouns and three verbs were then nearly evenly distributed to the “sentences,” forming 60 SOV sentences. Each sentence was completed with the particle /jo/, a common verb-final politeness marker in informal Korean.

The 60 sentences were elicited in two experimental sessions: the 30 sentences with “short” target words in Session A and the 30 sentences with “long” target words in Session B. The trials in each session were pseudo-randomized in such a way that trials from the same focus condition did not appear next to each other, and the focused constituent of a trial was not mentioned on its preceding trial.

2.4 Experimental procedure

The participants were tested individually upon being given consent by their parents in Hanyang Institute for Phonetics and Cognitive Sciences of Language at Hanyang University. A female native speaker of Seoul Korean administered the experiment after having received intensive training on how to conduct the experiment following the protocol. The experiment lasted about 60 min, including a short chat between the experimenter and the child before the first experimental session, and a short break between the two sessions. Audio recordings were made for each child in each session with a sampling rate of 44.1 kHz with 16-bit resolution. Video recordings were also made for some of the children for training purposes.

2.5 Prosodic annotation

The audio recordings from the participants were first orthographically annotated in Praat ([Boersma and Weenink, 2013](#)). Then, usable sentences were selected (1,602 or 83% from all the participants in the four age groups; 64% from the 4- to 5-year-olds, 70% from the 7- to 8-year-olds, 80% from the 10- to 11-year-olds, and 91% from the adults), and unusable ones were excluded from further analysis. A target sentence was considered unusable in any of the following cases: (1) the participant produced the target sentence before the experimenter asked the question, (2) the experimenter asked a different question than the intended question on that trial, (3) the experimenter did not provide an adequate description of the picture before she asked a question, (4) the sentence was produced with strong background noise, (5) the sentence was produced with word insertion, deletion or replacement, (6) the sentence was produced with self-repair or clearly perceivable hesitation, or (7) the sentence was produced with perceivable irregular voice quality or intonation caused by cold or unstable emotion.

The usable sentences were then annotated for phrasing, following the Korean Tones and Break Indices (K-ToBI) transcription conventions ([Jun, 2000, 2005](#)). That is, the boundaries immediately before the target words (i.e., between the subject-noun and object-noun) and after the target words (i.e., between the object-noun and verb) were annotated as an AP boundary, an IP boundary, or a phrase-internal phonological word boundary (hereafter PW boundary) by combining auditory impression and close inspection of prosodic cues to phrasing (e.g., tonal

⁶ There are two nominative case markers (i.e., /ka/ and /i/) and two accusative case markers (/il/ and /lil/) in Korean. The choice between each two variants depends on the presence of a coda consonant in the preceding noun.

TABLE 1 Words that occurred in the SOV sentences.

Subjects	개가 /kɛka/ “dog”	쥐가 /tʃwika/ “rat”	곰이 /komi/ “bear”	말이 /mali/ “horse”	소가 /soka/ “cow”	
Short objects	발을 /palil/ “foot”	비를 /pilil/ “rain”	불을 /pulil/ “fire”	팔을 /pʰalil/ “arm”	빵을 /p*anil/ “bread”	알을 /alil/ “egg”
Long objects	가방들을 /kapanʔilil/ “bags”	기둥들을 /kituʔtilil/ “pillars”	구두들을 /kututilil/ “shoes”	카드들을 /kʰatitilil/ “cards”	까마귀를 /k*amakwilil/ “crow”	안경들을 /ankʌŋʔilil/ “pairs of glasses”
Verbs	그려 /kiljʌ/ “draw”	만져 /mantʃʌ/ “touch”	찾아 /tʰatʃa/ “look for”			

patterns, boundary tones and breaks). The AP boundary is “a minimal phrasal disjuncture, with no strong subjective sense of pause” and is associated with AP tonal patterns as described in K-ToBI (Jun, 2005, p. 219). Word-final (i.e., pre-boundary) high tone and word-initial (i.e., post-boundary) low tone are taken as the typical AP boundary markers. The absence of voicing in word-initial lenis stops is also an informative indicator of an AP boundary (e.g., Jun, 1993; Cho et al., 2002). Moreover, word-initial (or post-boundary) strengthening in terms of perceptual clarity in the initial segment or syllable may also indicate an AP boundary, unless this cue contradicts another important cue such as a tonal cue. The IP boundary refers to phrasal boundaries that are demarcated by boundary tones and “a strong phrasal disjuncture, with a strong subjective sense of pause,” that is, either an “objective visible pause” or a “virtual pause” cued by final lengthening, as described in K-ToBI (Jun, 2005, p. 219). The PW boundary refers to word boundaries that are not demarcated by perceivable prosodic disjunctures in K-ToBI. It is worth noting that dephrasing does not consistently occur immediately after a focused word in Korean. In such a case, the AP boundary between the focused word and the post-focal word remains, but the pitch span of the post-focal word can be reduced (e.g., Jun and Lee, 1998; Kim et al., 2006; Lee and Xu, 2010). In the present study, when a boundary displayed the above-mentioned features of an AP boundary, but the post-boundary pitch span was noticeably reduced as compared to the pre-boundary pitch span, we annotated this boundary as an AP boundary instead of a PW boundary.

We conducted three rounds of annotation to maximize the reliability and agreement of the annotation. In the first round, the usable sentences were annotated by one transcriber (the first author) without access to information on the experimental conditions, following the above-described K-ToBI conventions, while a portion of the sentences (i.e., 81 sentences produced by two randomly selected participants) were jointly transcribed by two expert K-ToBI transcribers who were native speakers of Korean (the second and third authors), again without access to information on the experimental conditions. The two expert transcribers reached full agreement on the transcription of the 81 sentences. The Cohen's Kappa test on the annotation of the first transcriber and the expert transcribers for the 81 sentences revealed a very good inter-rater agreement for the boundaries before the target words ($K=0.811$, $p<0.0005$), and a good inter-rater agreement after the target words ($K=0.644$, $p<0.0005$) (Landis and Koch, 1977). The cases of disagreement were primarily concerned with the distinction between the AP boundary and the PW boundary. In the

second round, the first transcriber and the expert transcribers discussed the cases of disagreement, and agreed that the first transcriber should give more weight to three of the cues in her decision on AP and PW boundaries; namely, the word-initial and word-final tones and word-initial strengthening. The first transcriber then re-annotated all the usable sentences without access to the first-round annotation. The expert transcribers then jointly transcribed 23% of the usable sentences (i.e., 10 sentences randomly selected from each participant). The two expert transcribers reached full agreement on the transcription of the 23% sentences. The Cohen's Kappa test on the second-round annotation for 23% of the usable sentences revealed a very good inter-rater agreement between the first transcriber and the expert transcribers for the boundaries before the target words ($K=0.924$, $p<0.0005$), and after the target words ($K=0.897$, $p<0.0005$) (Landis and Koch, 1977). To reach a final agreement, a third-round annotation was conducted without access to the previous two rounds of annotation by two additional K-ToBI transcribers, who were native speakers of Korean and did not participate in previous rounds of annotation. The two expert transcribers subsequently examined the ambiguous ones reported by the two additional transcribers, and reached a full agreement on each boundary. The first transcriber then went through the third-round annotation and reached a final agreement with the expert transcribers and the additional transcribers. In this paper, we present an analysis of the data based on the third-round annotation.

3 Statistical analyses and results

Having annotated the prosodic boundaries before and after the sentence-medial target words in the sentences, we found that a large proportion of the sentences were produced as three separate APs in all age groups (57.8%⁷ for the 4- to 5-year-olds; 65.7% for the 7- to 8-year-olds; 62.2% for the 10- to 11-year-olds; 46.7% for the adults).

To statistically examine whether and how the children's use of phrasal boundaries before and after the target words may differ across focus conditions and age groups, we conducted mixed-effects multinomial logistic regression analyses using R Statistical Software

⁷ The percentages provided in the parentheses were calculated based on raw data.

(R Core Team, 2015) and the package Brms (Bürkner, 2017, 2018, 2021). Brms adopts a Bayesian approach with the Markov chain Monte Carlo (MCMC) method.

The random factors were SPEAKER (i.e., the participants) and SENTENCE (i.e., the target sentences). The dependent variable was BOUNDARY with three categories: AP boundary (reference category), IP boundary, and PW boundary. The independent variables (or fixed effects) were FOCUS and AGE. FOCUS referred to the focus conditions. For each analysis, we compared narrow focus to another focus condition to address a specific research question, so FOCUS always had two categories. AGE referred to the four age groups, with the adult group set as the reference category.

Three models were built using the aforementioned factors. Starting from an “empty” model (or Model 0) containing only the random factors, we added the effects of FOCUS and AGE to form Model 1, following Struiksma et al. (2022). The interaction between FOCUS and AGE was then added, forming Model 2. The method leave-one-out cross-validation (LOO) was used to evaluate model fit (Vehtari et al., 2017). The model with the lowest estimated looic was regarded as the best-fit model. The boundaries before and after the target words were analyzed separately.

As the model summary of the best-fit model does not straightforwardly show the difference between two focus conditions, or the difference between two focus conditions in each age group in the use of prosodic boundaries, we did follow-up analysis to answer the research questions. When the best-fit model was Model 2, containing the two-way interaction of focus and AGE, we examined the main effect of FOCUS in each age group, in order to address whether and how the speakers in each age group used prosodic boundaries to distinguish two focus conditions. When the best-fit model was Model 1 containing the main effects of FOCUS and AGE, we built and summarized a variant of model 1 containing only FOCUS as the fixed factor, in order to address how the speakers varied prosodic boundaries to distinguish two focus conditions, regardless of age. For concision, we report the co-efficient (B) and odds ratio ($\text{Exp}(B)$) from the models in the text; for transparency, we report summaries of these models and the best-fit models in Supplementary Tables 1–12.

3.1 Narrow focus vs. pre-focus

Analyzing the boundaries before the target words in the narrow focus and pre-focus conditions (Figure 2), we found that the best-fit model was Model 1 containing the effects of FOCUS and AGE (looic = 861.1) (Supplementary Table 1). FOCUS thus had a similar effect on the use of boundaries before the target words across age groups. A summary of the model containing only FOCUS (Supplementary Table 2) showed that the odds of the boundary before the target word being an IP boundary rather than an AP boundary in the narrow focus condition were 3.13 times as high as in the pre-focus condition ($B = -1.13$, $\text{Exp}(B) = 0.32$). The odds of the boundary before the target word being an PW boundary rather than an AP boundary in the pre-focus condition were 2.53 times as high as in the narrow focus condition ($B = 0.93$, $\text{Exp}(B) = 2.53$). In other words, the speakers were more likely to use an IP boundary, but less likely to use a PW boundary before the target word in the narrow focus condition than in the pre-focus condition, regardless of age.

Analyzing the boundaries after the target words, we found that the best-fit model was Model 2, containing the interaction of FOCUS and AGE (looic = 578.3) (Supplementary Table 3). Subsequent analysis on each age group showed that the model containing FOCUS was the best-fit model for each age group (looic = 242.3 for the adults; looic = 84.3 for the 4- to 5-year-olds; looic = 111.1 for the 7- to 8-year-olds; looic = 143.0 for the 10- to 11-year-olds) (Supplementary Tables 4–7).

For the adults, the odds of the boundary after the target word being an IP boundary rather than an AP boundary in the pre-focus condition were 1.89 times as high as in the narrow focus condition ($B = 0.64$, $\text{Exp}(B) = 1.89$). The odds of the boundary after the target word being an PW boundary rather than an AP boundary in the narrow focus condition were 110559.84 times as high as in the pre-focus condition ($B = -11.61$, $\text{Exp}(B) = 0.00$). In other words, the adults were less likely to use an IP boundary, but more likely to use a PW boundary after the target word in the narrow focus condition than in the pre-focus condition.

For the 4- to 5-year-olds, the odds of the boundary after the target word being an IP boundary rather than an AP boundary in the pre-focus condition were 10.30 times as high as in the narrow focus condition ($B = 2.33$, $\text{Exp}(B) = 10.30$). The odds of the boundary after the target word being an PW boundary rather than an AP boundary in the narrow focus condition were 7131.28 times as high as in the pre-focus condition ($B = -8.87$, $\text{Exp}(B) = 0.00$). In other words, the 4- to 5-year-olds were less likely to use an IP boundary, but more likely to use a PW boundary after the target word in the narrow focus condition than in the pre-focus condition.

For the 7- to 8-year-olds, the odds of the boundary after the target word being an IP boundary rather than an AP boundary in the pre-focus condition were 28984.8 times as high as in the narrow focus condition ($B = 10.27$, $\text{Exp}(B) = 28984.8$). The odds of the boundary after the target word being an PW boundary rather than an AP boundary in the narrow focus condition were 33.33 times as high as in the pre-focus condition ($B = -3.50$, $\text{Exp}(B) = 0.03$). In other words, the 7- to 8-year-olds were less likely to use an IP boundary, but more likely to use a PW boundary after the target word in the narrow focus condition than in the pre-focus condition.

For the 10- to 11-year-olds, the odds of the boundary after the target word being an IP boundary rather than an AP boundary in the pre-focus condition were 9.79 times as high as in the narrow focus condition ($B = 2.28$, $\text{Exp}(B) = 9.79$). The odds of the boundary after the target word being an PW boundary rather than an AP boundary in the narrow focus condition were 64533.95 times as high as in the pre-focus condition ($B = -11.07$, $\text{Exp}(B) = 0.00$). In other words, the 10- to 11-year-olds were less likely to use an IP boundary, but more likely to use a PW boundary after the target word in the narrow focus condition than in the pre-focus condition.

As an interim summary, the children used IP boundaries more frequently but used PW boundaries less frequently before the target words in the narrow focus condition than in the pre-focus condition, similar to the adults. These results indicated that the children preferred inserting a large prosodic boundary (IP) immediately before the word in the narrow focus condition, similar to the adults; they also preferred deleting the boundary between the two pre-focal words and producing them as one larger AP or IP when the sentence-final verb was focused, similar to the adults.

As for the boundaries after the target words, the speakers in all age groups used IP boundaries less frequently but used PW boundaries

more frequently in the narrow focus condition than in the pre-focus condition. The results indicated that they frequently dephrased the post-focal word in the narrow focus condition, and frequently inserted a large prosodic boundary (IP) immediately before the focused sentence-final word in the pre-focus condition. However, the children differed from the adults in absolute frequency in their use of prosodic boundaries after the target words. The 7- to 8-year-olds were least similar to the adults. To distinguish the narrow focus and pre-focus conditions, while the other age groups seemed to rely more on the use of PW boundaries (or post-focus dephrasing) rather than the use of IP boundaries after the target words, the 7- to 8-year-olds relied more on the use of IP boundaries than the use of PW boundaries after the target words. The 10- to 11-year-olds were most similar to the adults in their use of prosodic boundaries in terms of absolute frequency.

3.2 Narrow focus vs. post-focus

Analyzing the boundaries before the target words in the narrow focus and post-focus conditions, we found that the best-fit model was Model 0 containing only the random effects ($\text{looic}=863.2$), indicating that the speakers did not vary the boundaries before the target words to distinguish narrow focus from post-focus, regardless of age.

Analyzing the boundaries after the target words, we found that the best-fit model was Model 0 containing only the random effects ($\text{looic}=659.7$), indicating that the speakers did not vary the boundaries after the target words to distinguish narrow focus from post-focus, regardless of age.

As an interim summary, the children did not vary the boundaries before or after the target words to distinguish narrow focus from post-focus, similar to the adults. As the speakers tended to insert an AP or IP boundary before the sentence-medial target word and dephrase its following word when the target word was under narrow focus, this part of the results indicated that they also did so when the target word was post-focus.

3.3 Narrow focus vs. broad focus

Analyzing the boundaries before the target words in the narrow focus and broad focus conditions (Figure 3), we found that the best-fit model was Model 2, containing the interaction of FOCUS and AGE ($\text{looic}=782.3$) (Supplementary Table 8). Subsequent analysis on each age group showed that the model containing FOCUS was the best-fit model for the 10- to 11-year-olds ($\text{looic}=205.8$) and the adults ($\text{looic}=258.6$) (Supplementary Tables 9–10), but the model containing only the random effects was the best-fit model for the 4- to 5-year-olds ($\text{looic}=154.6$) and 7- to 8-year-olds ($\text{looic}=165.9$). In other words, the 10- to 11-year-olds varied prosodic boundaries before the target words to distinguish narrow focus from broad focus, similar to the adults, but the 4- to 5-year-olds and 7- to 8-year-olds did not.

For the adults, the odds of the boundary before the target word being an IP boundary rather than an AP boundary in the broad focus condition were 4.08 times as high as in the narrow focus condition ($B=1.41$, $\text{Exp}(B)=4.08$). The odds of the boundary before the target word being a PW boundary rather than an AP boundary in the narrow focus condition were 26041.80 times as high as in the broad focus condition ($B=-10.17$, $\text{Exp}(B)=0.00$). In other words, the adults

were more likely to use an IP boundary, but less likely to use a PW boundary before the target word in the broad focus condition than in the narrow focus condition.

For the 10- to 11-year-olds, the odds of the boundary before the target word being an IP boundary rather than an AP boundary in the broad focus condition were 4.31 times as high as in the narrow focus condition ($B=1.46$, $\text{Exp}(B)=4.31$). The odds of the boundary before the target word being a PW boundary rather than an AP boundary in the narrow focus condition were 3.13 times as high as in the broad condition ($B=-1.15$, $\text{Exp}(B)=0.32$). In other words, the 10- to 11-year-olds were more likely to use an IP boundary, but less likely to use a PW boundary before the target word in the broad focus condition than in the narrow focus condition, similar to the adults.

Analyzing the boundaries after the target words, we found that the best-fit model was Model 1 containing the effects of FOCUS and AGE ($\text{looic}=670.1$) (Supplementary Table 11). FOCUS thus had a similar effect on the use of boundaries after the target words across age groups. A summary of the model containing only FOCUS (Supplementary Table 12) showed that the odds of the boundary after the target word being an IP boundary rather than an AP boundary in the broad focus condition were 5.58 times as high as in the narrow focus condition ($B=1.72$, $\text{Exp}(B)=5.58$). The odds of the boundary after the target word being a PW boundary rather than an AP boundary in the narrow focus condition were 5.56 times as high as in the broad focus condition ($B=-1.70$, $\text{Exp}(B)=0.18$). In other words, the three groups of children used IP boundaries more frequently but used PW boundaries less frequently after the focal words in the broad focus condition than in the narrow focus condition, similar to the adults.

As an interim summary, the adults used IP boundaries more frequently before and after the target words in the broad focus condition than in the narrow focus condition, indicating their preference for producing the words as separate IPs when the sentence was in broad focus. The adults used PW boundaries more frequently after the target words in the narrow focus condition than in the broad focus condition, indicating their preference for post-focus dephrasing in the narrow focus condition. The 10- to 11-year-olds were adult-like in distinguishing narrow focus from broad focus. However, the two younger groups of children only varied the boundaries after the target words in the same way as the adults did.

3.4 Narrow focus vs. contrastive focus

Analyzing the boundaries before the target words in the narrow focus and contrastive focus conditions, we found that the best-fit model was Model 0 containing only the random effects ($\text{looic}=854.6$), indicating that the speakers did not vary the boundaries before the target words to distinguish narrow focus from contrastive focus, regardless of age.

Analyzing the boundaries after the target words, we found that the best-fit model was Model 0 containing only the random effects ($\text{looic}=716.7$), indicating that the speakers did not vary the boundaries after the target words to distinguish narrow focus from contrastive focus, regardless of age.

As an interim summary, the children did not vary the boundaries before or after the target words to distinguish narrow focus from contrastive focus, similar to the adults. In other words, the children and adults marked contrastive focus similarly to narrow focus via prosodic phrasing.

4 General discussion

To further the current understanding of the effect of transparency of phonological form-meaning mapping on the acquisition of prosodic focus marking, we have examined how Korean-speaking 4- to 5-year-olds, 7- to 8-year-olds, and 10- to 11-year-olds varied prosodic boundaries to distinguish narrow focus from non-focus (i.e., pre-focus and post-focus) and two other types of focus (i.e., broad focus and contrastive focus) in semi-spontaneous production of SOV sentences, as compared to adults.

Regarding the prosodic realization of narrow focus as compared to non-focus, we first compared narrow focus to pre-focus and found that the children in all age groups preferred inserting a large prosodic boundary (i.e., IP boundary) immediately before the focused target word in the narrow focus condition, similar to the adults, and in line with previous findings on Korean-speaking adults' read speech (e.g., Jun, 1993; Jun and Kim, 2007; Jeon and Nolan, 2017). They also preferred deleting the boundary between the pre-focal words and producing them as one larger AP or IP when the sentence-final verb was focused, similar to the adults. This observation coincides with an earlier finding that pre-focal words tend to be dephrased in adults' production of read speech (Oh, 1999). Regarding the use of prosodic boundaries after the target words, the children in all age groups frequently used a PW boundary (i.e., dephrasing the post-focal word) in the narrow focus condition, and frequently inserted a large prosodic boundary (IP) immediately before the focused sentence-final word in the pre-focus condition, similar to the adults. However, while the other age groups seemed to rely more on the use of PW boundaries (or post-focus dephrasing) rather than the use of IP boundaries after the target words, the 7- to 8-year-olds relied more on the use of IP boundaries than the use of PW boundaries after the target words, and thus were least similar to the adults. The 10- to 11-year-olds were most similar to the adults in their use of prosodic boundaries in terms of absolute frequency. Regarding the comparison between narrow focus and post-focus, we found that while the children tended to use an AP or IP boundary before the sentence-medial target word and dephrase its following word when the target word was under narrow focus, they also did so when the target word was post-focus, similar to the adults.

Thus, our prediction that Korean-speaking children will be fully adult-like at the age of 7–8 years or later in more frequently using an AP and/or IP boundary before a word and a PW boundary after it in the narrow focus condition than in the non-focus conditions (i.e., pre-focus or post-focus) is only partly borne out. We have unexpectedly observed a protracted developmental path in Korean-speaking children, as the 7- to 8-year-olds relied more on the use of large prosodic boundaries (i.e., IP boundaries) than PW boundaries (related to post-focus dephrasing) to distinguish the narrow focus and pre-focus conditions, different from the other age groups. We suggest two speculations for this finding. First, as the 7- to 8-year-olds start to take read-aloud practice in primary school, their preference to large prosodic boundaries might be from the influence of read speech. However, as previous studies on adults' read speech usually analyzed different focus conditions separately, we lack comparable findings on how prosodic boundaries are used to distinguish focus conditions in read speech. Moreover, we did not observe a similar pattern in the 10- to 11-year-olds, who had an equal chance of being influenced by read speech in class. Second, the 7- to 8-year-olds might have been more engaged in the picture-matching game, and thus put in more

effort in providing answers to the experimenter, compared to the 10- to 11-year-olds and the adults. We speculate that in an edge-prominence language like Korean, more effort in marking focus might lead to more frequent use of prosodic phrasing related to large prosodic boundaries in distinguishing focus conditions. The game was designed in such a way that it would suit the youngest children in the study. It is thus plausible that the oldest children and the adults did not put in more effort in production than needed in encoding focus. A similar observation about 10- to 11-year-olds engaging less with the game and making less effort to answer questions than younger children was reported in Romøren and Chen (2021).

As for the distinction between narrow focus and broad focus, only the 10- to 11-year-olds were fully adult-like. They preferred producing the words as separate IPs when the sentence was in broad focus, and preferred post-focus dephrasing in the narrow focus condition. However, the two younger groups of children only varied the boundaries after the target words in the same way as the adults did to distinguish the two focus conditions. These results largely support our prediction concerning the effect of focal constituent size.

Regarding the comparison between narrow focus and contrastive focus, we did not find any evidence of the speakers distinguishing the two types of focus using prosodic boundaries, regardless of age. The results fully support our prediction regarding the effect of contrastivity.

In previous studies on the acquisition of phonological focus marking, Dutch-speaking children used accentuation close to the ceiling across narrow focus, broad focus, and contrastive focus conditions, showing no evidence of using accent placement in distinguishing focus types, similar to Dutch-speaking adults (Romøren, 2016), although the exact accent types used for marking the three types of focus were not reported in Romøren (2016). Swedish-speaking children used the prominence-marking high tone more frequently in narrow focus than in broad focus, but did not distinguish narrow focus from contrastive focus using this cue, similar to Swedish-speaking adults (Romøren, 2016). Thus, the current findings, along with those from previous studies, imply that languages differ both in whether and how focus types are distinguished by language-specific phonological cues, and in the acquisition of phonological marking of these focus types.

Based on the findings of the present study, we can depict Korean-speaking children's developmental path of phonological focus marking from the age of 4 to 11 as follows: At the age of 4 to 5, Korean-speaking children use prosodic phrasing to mark focus and distinguish different types of focus in a largely adult-like manner, though their use of prosodic boundaries for focus-marking purposes is not fully adult-like in terms of absolute frequency. In contrast, 4- to 5-year-old Swedish-speaking children were fully adult-like in the use of the prominence-marking high tone in sentence-final position, and largely adult-like in sentence-medial position in terms of manner and frequency (Romøren, 2016; Romøren and Chen, 2021). The difference between Korean-speaking and Swedish-speaking children is further evident at later stages. For example, at the age of 7 to 8, Korean-speaking children tend to rely more on the use of large prosodic boundaries (i.e., IP boundaries) than PW boundaries (or dephrasing) when distinguishing narrow focus from pre-focus, different from adults. At the age of 10 to 11, they exhibit fully adult-like abilities in distinguishing focus conditions. The results in general support our hypothesis regarding Korean-speaking children's rate of acquisition of phonological focus marking, compared to that of Swedish-speaking children and Dutch-speaking children.

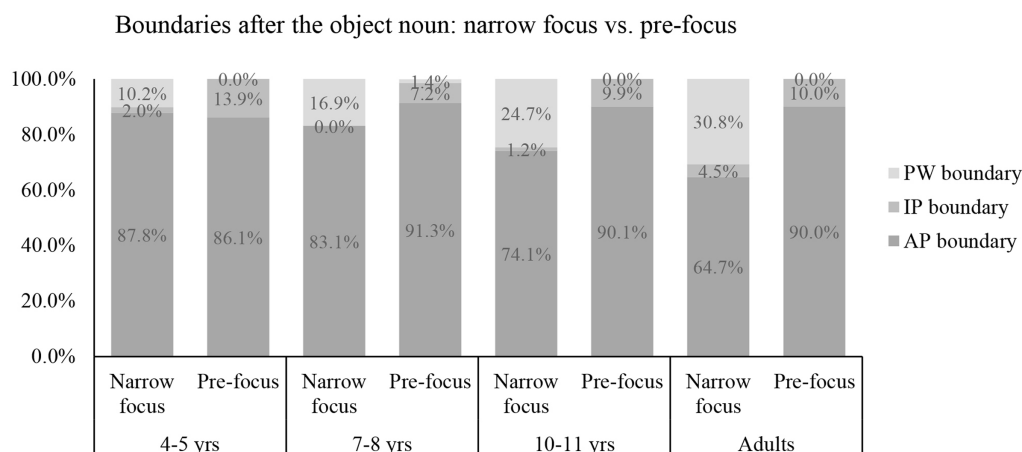


FIGURE 2

The use of boundaries after the object-noun in the narrow focus condition and pre-focus condition.

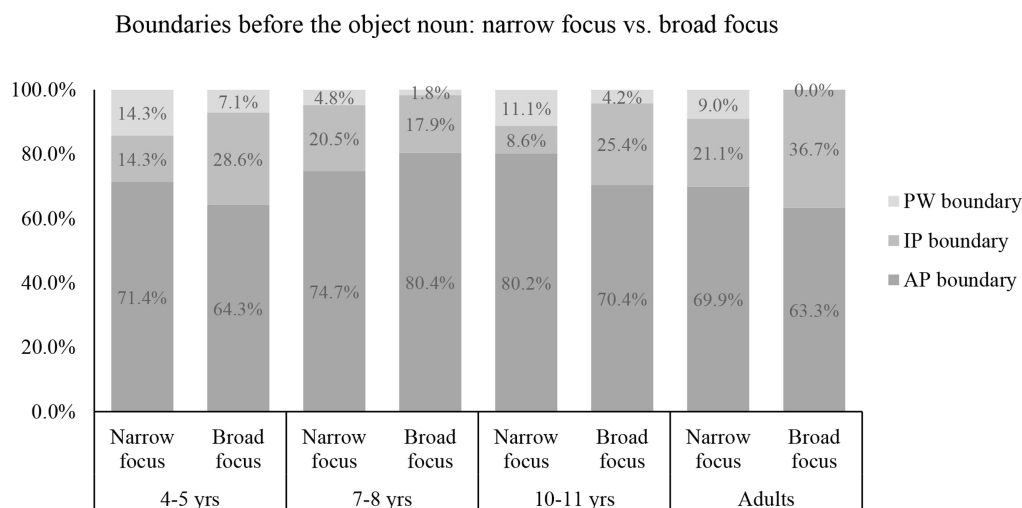


FIGURE 3

The use of boundaries before the object-noun in the narrow focus condition and broad focus condition.

The results have further implications for understanding cross-linguistic variation in the acquisition of focus marking. As discussed at the outset of the paper, Swedish employs a highly transparent phonological means of focus marking, so that words under narrow focus are consistently assigned a word-final high tone (Bruce, 2007; Romøren, 2016). In other words, Swedish demonstrates a clear and direct mapping between the phonological form and the focus-related meaning. On the other hand, while focus marking in Korean is typically achieved through phrasing (involving the initiation of a large prosodic constituent such as an IP) (e.g., Jun, 1993, 1998, 2000, 2005), phrasing is not exclusively used for focus marking, indicating a less transparent form-meaning mapping compared to that observed in Swedish. Given these cross-linguistic differences between Swedish and Korean, our results lend support to Chen's (2018) view that the transparency of the form-meaning mapping between phonological cues and focus conditions influences the rate of acquisition in prosodic focus marking across languages.

Let us now compare our Korean results with those observed in Dutch-speaking children. Recall that Dutch-speaking 4- to 5-year-olds were not fully adult-like in their choice of accent type for focus-marking purposes (Chen, 2011b). In contrast, at these ages (4 to 5 years), Korean-speaking children exhibited some phonological focus marking patterns that were more adult-like than their Dutch-speaking peers. Thus, our results suggest that Korean-speaking children tend to acquire an adult-like way of focus marking relatively earlier than their Dutch-speaking peers. However, the difference between Korean-speaking and Dutch-speaking children does not seem to be fully in line with the effect of transparency based on consistency in the association between a form and a meaning. This is because there is no apparent difference in the consistency of the form-meaning mapping between Korean and Dutch. In other words, phrasing and pitch accent, which are used for focus marking in each language, respectively, can also occur in various other non-focused contexts. However, our results are compatible with a transparency hypothesis based on perceptual

transparency between a form and a meaning. That is, if we extend the notion of the degree of transparency in the form-meaning mapping to include perceptual transparency of the cues to focus, the earlier acquisition of focus marking by Korean-speaking children can still be understood as a reflection of cross-linguistic differences in the transparency of the form-meaning mapping.

5 Conclusion and limitations

In conclusion, our findings on Korean-speaking children support Chen's (2018) view that a higher degree of transparency in the form-meaning mapping between phonological cues and focus leads to a faster rate of acquisition in prosodic focus marking. Further, we demonstrate that a greater diversity in phonological forms, such as the use of phrasing for purposes other than focus marking in Korean, can slow down the rate of acquisition. Thus, our study not only provides new experimental evidence for the role of transparency in form-meaning mapping as a determinant of children's acquisition of focus marking, but also expands our current understanding of the notion of the degree of transparency in the form-meaning mapping to include perceptual salience of the phonological cue.

Some important questions remain to be addressed in future studies. For example, we did not observe evidence of Korean-speaking adults and children distinguishing narrow focus from post-focus as well as narrow focus from contrastive focus using prosodic phrasing. Given that these focus conditions are distinguished phonologically and/or phonetically in many other languages, they may be distinguished with pitch- and duration-related phonetic cues as well as segmental cues in Korean. It will be insightful to study whether and how other phonetic focus-marking cues are used in Korean when the primary cue, prosodic phrasing, does not suffice to distinguish two focus conditions (but see Cho et al., 2011; Hatcher et al., 2023, for related data), as well as how Korean-speaking children acquire the use of the phonetic cues. Moreover, as Korean utilizes both the left and right edges of prosodic phrases for phrasing and focus marking, another interesting question for future research is whether our findings on the acquisition of prosodic focus marking in Korean-speaking children can be generalized to children acquiring other edge-prominence languages, like Mongolian, in which prosodic phrasing and focus are typically marked at the left edge of prosodic units (Karlsson, 2014). Thus, more studies under the same theoretical framework adopting similar experimental methods need to be done to broaden our understanding of the acquisition of prosodic focus marking.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical review and approval were not required for the study on human participants in accordance with the local legislation and institutional requirements at the time of testing. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

AY: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Resources, Validation, Visualization, Writing – original draft, Writing – review & editing. TC: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Resources, Validation, Writing – review & editing. SK: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Resources, Validation, Writing – review & editing. AC: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by a VIDI grant awarded to AC by the Dutch Research Council (NWO) (grant number 276-89-001). This work was also supported in part by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2023S1A5A2A03084924) to TC and SK, and by the Independent Innovation Fund of Tianjin University (2019XSC-0034; 2024XSC-0023) to AY.

Acknowledgments

We would like to thank Hanyang Institute for Phonetics and Cognitive Sciences of Language for its great help with the experiment. We express our special gratitude to the children and their parents for their participation.

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2024.1352280/full#supplementary-material>

References

- Baumann, S., Grice, M., and Steindamm, S. (2006). Prosodic marking of focus domains-categorical or gradient. *Proceedings of the International Conference on Speech Prosody 3*, 301–304. Dresden, Germany.
- Bishop, J., Kuo, G., and Kim, B. (2020). Phonology, phonetics, and signal-extrinsic factors in the perception of prosodic prominence: evidence from rapid prosody transcription. *J. Phon.* 82:100977. doi: 10.1016/j.wocn.2020.100977
- Boersma, P., and Weenink, D. (2013). Praat: doing phonetics by computer (Version 5.3) [Computer software]. Available at: <http://www.praat.org/>
- Bruce, G. (2007). “Components of a prosodic typology of Swedish intonation” in *Tones and tunes. typological and comparative studies in word and sentence prosody*. eds. T. Riad and C. Gussenhoven, vol. 1 (Berlin: Mouton de Gruyter), 113–146.
- Bürkner, P. (2017). Brms: an R package for Bayesian multilevel models using Stan. *J. Stat. Softw.* 80, 1–28. doi: 10.18637/jss.v080.i01
- Bürkner, P. (2018). Advanced Bayesian multilevel modeling with the R package brms. *R J.* 10, 395–411. doi: 10.32614/RJ-2018-017
- Bürkner, P. (2021). Bayesian item response modeling in R with brms and Stan. *J. Stat. Softw.* 100, 1–54. doi: 10.18637/jss.v100.i05
- Chen, A. (2009). “The phonetics of sentence-initial topic and focus in adult and child Dutch” in *Phonetics and phonology: Interactions and interrelations*. eds. M. Vigário, S. Frota and M. J. Freitas (Amsterdam: Benjamins), 91–106.
- Chen, A. (2011a). “The developmental path to phonological focus-marking in Dutch” in *Prosodic categories: Production, perception and comprehension*. eds. S. Frota, G. Elordieta and P. Prieto (Dordrecht: Springer), 93–109.
- Chen, A. (2011b). Tuning information packaging: intonational realization of topic and focus in child Dutch. *J. Child Lang.* 38, 1055–1083. doi: 10.1017/S0305000910000541
- Chen, A. (2018). “Get the focus right across languages: acquisition of prosodic focus-marking in production” in *Prosodic development*. eds. P. Prieto and Esteve-Gibert (Amsterdam: Benjamins).
- Chen, Y., and Braun, B. (2006). Prosodic realization of information structure categories in standard Chinese. *Proceedings of the international conference on speech prosody 3*, Dresden, Germany.
- Chen, A., Esteve-Gibert, N., Prieto, P., and Redford, M. A. (2020). “Development in phrase-level prosody from infancy to late childhood” in *The Oxford handbook of language prosody*. eds. C. Gussenhoven and A. Chen (Oxford University Press), 553–562.
- Cho, T. (2022). “The phonetics-prosody interface and prosodic strengthening in Korean” in *The Cambridge Handbook of Korean Linguistics*. Eds. S. Cho, J. Whitman (Cambridge: Cambridge University Press), 248–293.
- Cho, T., Jun, S. A., and Ladefoged, P. (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *J. Phon.* 30, 193–228. doi: 10.1006/jpho.2001.0153
- Cho, T., Lee, Y., and Kim, S. (2011). Communicatively driven versus prosodically driven hyper-articulation in Korean. *J. Phon.* 39, 344–361. doi: 10.1016/j.wocn.2011.02.005
- Christophe, A., Dupoux, E., Bertoncini, J., and Mehler, J. (1994). Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition. *J. Acoust. Soc. Am.* 95, 1570–1580. doi: 10.1121/1.408544
- Christophe, A., Mehler, J., and Sebastián-Gallés, N. (2001). Perception of prosodic boundary correlates by newborn infants. *Infancy* 2, 385–394. doi: 10.1207/S15327078IN0203_6
- Cole, J., Mahrt, T., and Roy, J. (2017). Crowd-sourcing prosodic annotation. *Comput. Speech Lang.* 45, 300–325. doi: 10.1016/j.csl.2017.02.008
- Cole, J., and Shattuck-Hufnagel, S. (2016). New methods for prosodic transcription: capturing variability as a source of information. *Lab. Phonol.* 7:8. doi: 10.5334/labphon.29
- Escudero, D., Aguilar, L., del Mar Vanrell, M., and Prieto, P. (2012). Analysis of inter-transcriber consistency in the Cat_ToBI prosodic labeling system. *Speech Comm.* 54, 566–582. doi: 10.1016/j.specom.2011.12.002
- Grice, M., Reyelt, M., Benzmueller, R., Mayer, J., and Batliner, A. (1996). Consistency in transcription and labelling of German intonation with GToBI. *Proceeding of Fourth International Conference on Spoken Language Processing*, Philadelphia. pp. 1716–1719.
- Grünloh, T., Lieven, E., and Tomasello, M. (2015). Young children’s intonational marking of new, given and contrastive referents. *Lang. Learn. Dev.* 11, 95–127. doi: 10.1080/15475441.2014.889530
- Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.
- Gussenhoven, C. (2007). “Types of focus in English” in *Topic and focus: cross-linguistic perspectives on meaning and intonation*. eds. C. Lee, M. Gordon and D. Büring (Heidelberg, New York, London: Springer), 83–100.
- Hanssen, J., Peters, J., and Gussenhoven, C. (2008). Prosodic effects of focus in Dutch declaratives. *Proceedings of the International Conference on Speech Prosody*, Campinas, Brazil. 4, 609–612.
- Hatcher, R., Joo, H., Kim, S., and Cho, T. (2023). “How does focus-induced prominence influence realization of edge tones and segmental anchoring in Seoul Korean?” in *Proceedings of the 20th international congress of phonetic sciences*. eds. R. Skarnitzl and J. Volin (Prague: Guarant International), 1484–1488.
- Holzgrefe-Lang, J., Wellmann, C., Höhle, B., and Wartenburger, I. (2018). Infants’ processing of prosodic cues: electrophysiological evidence for boundary perception beyond pause detection. *Lang. Speech* 61, 153–169. doi: 10.1177/0023830917730590
- Hornby, P. A., and Hass, W. A. (1970). Use of contrastive stress by preschool children. *J. Speech Hear. Res.* 13, 395–399. doi: 10.1044/jshr.1302.395
- Jeon, H. S., and Nolan, F. (2017). Prosodic marking of narrow focus in Seoul Korean. *Lab. Phonol.* 8, 1–30. doi: 10.5334/labphon.48
- Johnson, E. K., and Seidl, A. (2008). Clause segmentation by 6-month-old infants: a crosslinguistic perspective. *Infancy* 13, 440–455. doi: 10.1080/15250000802329321
- Jun, S. A. (1993). *The phonetics and phonology of Korean prosody* (Doctoral dissertation). Ohio: Ohio State University.
- Jun, S. A. (1998). The accentual phrase in the Korean prosodic hierarchy. *Phonology* 15, 189–226. doi: 10.1017/S0952675798003571
- Jun, S. A. (2000). K-ToBI (Korean ToBI) labeling conventions. *UCLA Working Pap. Phonet.* 99, 149–173.
- Jun, S. A. (2005). “Korean intonational phonology and prosodic transcription” in *Prosodic typology: the phonology of intonation and phrasing*. ed. S. A. Jun (Oxford: Oxford University Press), 201–229.
- Jun, S. A. (2007a). Phonological development of Korean: a case study. *UCLA Working Pap. Phonet.* 105, 51–65.
- Jun, S. A. (2007b). “The intermediate phrase in Korean intonation: evidence from sentence processing” in *Tones and Tunes: Studies in word and sentence prosody*. eds. C. Gussenhoven and T. Riad (Berlin: Mouton de Gruyter), 143–167.
- Jun, S. A. (2011). Prosodic markings of complex NP focus, syntax, and the pre-/post-focus string. *Proceedings of the west coast conference on formal linguistics 28*, 214–230. Somerville, MA: Cascadia Press.
- Jun, S. A. (2014). “Prosodic typology: by prominence type, word prosody, and macro-rhythm” in *Prosodic typology II: the phonology of intonation and phrasing*. ed. S. A. Jun, vol. 2014 (Oxford: Oxford University Press), 520–540.
- Jun, S. A., and Kim, H. S. (2007). VP focus and narrow focus in Korean. *Proceedings of the international congress of phonetic sciences*, Saarbrücken, Germany.
- Jun, S. A., Kim, H. S., Lee, H. J., and Kim, J. B. (2006). An experimental study on the effect of argument structure on VP focus. *Korean Linguist.* 13, 89–113. doi: 10.1075/kl.13.04saj
- Jun, S. A., and Lee, H. J. (1998). Phonetic and phonological markers of contrastive focus in Korean. *Proceedings of the International Conference on Spoken Language Processing*, 4, 1295–1298. Sydney, Australia.
- Jun, S. A., Lee, S., Kim, K., and Lee, Y. (2000). Labeler agreement in transcribing Korean intonation with K-ToBI. *Proc. ICSLP 3*, 211–214.
- Karlsson, A. M. (2014). *The intonational phonology of Mongolian*, in *Prosodic Typology II: The Phonology of Intonation and Phrasing*. Ed. Sun-Ah Jun, (Oxford, 2014; online edn, Oxford Academic, 1 Apr. 2014).
- Kim, H. S., Jun, S. A., Lee, H. J., and Kim, J. B. (2006). Argument structure and focus projection in Korean. *Proceedings of speech prosody 3*, Dresden, Germany.
- Kim, J. J., Kim, S., and Cho, T. (2024). Preboundary lengthening and articulatory strengthening in Korean as an edge-prominence language. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 15, 1–36. doi: 10.16995/labphon.9880
- Kügler, F., and Calhoun, S. (2021). “Prosodic encoding of information structure: a typological perspective” in *The Oxford handbook of language prosody*. eds. C. Gussenhoven and A. Chen (Oxford: Oxford University Press), 454–467.
- Landis, J. R., and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174.
- Lee, Y. C., and Xu, Y. (2010). Phonetic realization of contrastive focus in Korean. *Proc. Speech Prosody 5*, 1–4.
- Liu, L. Z. H. (2017). *The development of prosodic focus-marking in early bilinguals’ L2: a study of Bai-mandarin early bilinguals’ mandarin*. Utrecht: LOT.
- MacWhinney, B., and Bates, E. (1978). Sentential devices for conveying givenness and newness: a cross- cultural developmental study. *J. Verbal Learn. Verbal Behav.* 17, 539–558. doi: 10.1016/S0022-5371(78)90326-2
- Oh, M. (1999). Korean prosodic structure and focus. *Proceedings of the International Conference of Phonetic Sciences: San Francisco 14*, 1517–1520.
- R Core Team (2015). *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

- Romøren, A. S. H. (2016). *Hunting highs and lows: the acquisition of prosodic focus marking in Swedish and Dutch*. Utrecht: LOT.
- Romøren, A. S. H., and Chen, A. (2021). The acquisition of prosodic marking of narrow focus in central Swedish. *J. Child Lang.* 49, 213–238. doi: 10.1017/S0305000920000847
- Seidl, A. (2007). Infants' use and weighting of prosodic cues in clause segmentation. *J. Mem. Lang.* 57, 24–48. doi: 10.1016/j.jml.2006.10.004
- Seidl, A., and Cristià, A. (2008). Developmental changes in the weighting of prosodic cues. *Dev. Sci.* 11, 596–606. doi: 10.1111/j.1467-7687.2008.00704.x
- Shih, C. (1988). Tone and intonation in mandarin. *Working papers of the Cornell phonetics. Laboratory* 3, 83–109.
- Struiksma, M. E., De Mulder, H. N., and Van Berkum, J. J. (2022). Do people get used to insulting language? *Front. Commun.* 7:910023. doi: 10.3389/fcomm.2022.910023
- Thorson, J., and Morgan, J. (2020). Prosodic realization of NEW, GIVEN, and CORRECTIVE referents in the spontaneous speech of toddlers. *J. Child Lang.* 47, 1–28. doi: 10.1017/S0305000920000434
- Vallduví, E., and Engdahl, E. (1996). The linguistic realization of information packaging. *Linguistics* 34, 459–520. doi: 10.1515/ling.1996.34.3.459
- Vehtari, A., Gelman, A., and Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat. Comput.* 27, 1413–1432. doi: 10.1007/s11222-016-9696-4
- Wells, B., Peppé, S., and Goulondris, N. (2004). Intonation development from five to thirteen. *J. Child Lang.* 31, 749–778. doi: 10.1017/S030500090400652X
- Wu, W. L., and Xu, Y. (2010). Prosodic focus in Hong Kong Cantonese without post-focus compression. *Proc. Speech Prosody* 5, 1–4.
- Xu, Y. (1999). Effects of tone and focus on the formation and alignment of f0 contours. *J. Phon.* 27, 55–105. doi: 10.1006/jpho.1999.0086
- Yang, A., and Chen, A. (2018). The developmental path to adult-like prosodic focus-marking in mandarin Chinese-speaking children. *First Lang.* 38, 26–46. doi: 10.1177/0142723717733920



OPEN ACCESS

EDITED BY

Bradley Hoot,
DePaul University, United States

REVIEWED BY

Laura Dominguez,
University of Southampton, United Kingdom
Julio César López Otero,
University of Houston, United States

*CORRESPONDENCE

Liz Smeets
✉ lsmeets@yorku.ca

RECEIVED 12 March 2024

ACCEPTED 23 July 2024

PUBLISHED 13 September 2024

CITATION

Smeets L (2024) L1 grammatical attrition
through the acquisition of competing L2
discourse features.
Front. Psychol. 15:1399870.
doi: 10.3389/fpsyg.2024.1399870

COPYRIGHT

© 2024 Smeets. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License](#)
(CC BY). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

L1 grammatical attrition through the acquisition of competing L2 discourse features

Liz Smeets*

Department of Languages, Literatures and Linguistics, York University, Toronto, ON, Canada

A question in language acquisition research is whether attrition can affect L1 grammatical representation, and if so, under what conditions. This paper tests the Attrition via Acquisition (AvA) model, which takes a Feature Reassembly approach to predict how, in case on high degrees of similarity between the L1 and L2, the acquisition of L2 discourse-driven morpho-syntactic properties may affect L1 feature representations after a prolonged change in the speaker's primary linguistic input during adulthood. As a test case, we use the different features (*specificity* versus *discourse anaphoricity*) associated with Clitic Left Dislocation (CLLD) in Romanian and Italian, examining the grammars of Romanian first-generation immigrants with either L2 Italian or L2 English (a language without CLLD). Using a context-dependent Acceptability Judgment task and a Written Elicitation task we found evidence for L2-induced grammatical attrition, resulting in the addition of an L2 option without the loss of an L1 option, as predicted by the AvA. Attrition was found for participants who immigrated during adolescence or early adulthood and who are more likely to consider Italian their most proficient and most used language. We compare our findings on attrited L1 grammars to the results of a recent study reporting on near-native L2 Italian and L2 Romanian grammars by Romanian and Italian native speakers. Our findings contribute to an increasing body of literature showing that L1 attriters and L2 learners can end up with very similar grammars and confirm the importance of studying second language acquisition and L1 loss within a broader picture of bilingual development.

KEYWORDS

clitic left dislocation, L1 attrition, L2 transfer, Italian, Romanian, discourse-syntax interface

1 Introduction

The current study focuses on the potential modification or restructuring of a speaker's L1 grammatical representation ('grammatical attrition') in individuals whose primary linguistic input has changed for a prolonged period of time due to immigration to a country where a different language is spoken. The available empirical evidence on grammatical attrition largely supports the idea that effects of attrition are relatively minor in late bilinguals compared to early bilinguals such as heritage speakers; an overview by Montrul (2008) reports that studies on adult L1 attrition of morpho-syntactic phenomena rarely ever find that attriters make morpho-syntactic errors in more than 5% of the contexts in which a specific morpheme is required. Neurolinguistic research also supports the idea that potential attriters behave like native speakers on morpho-syntactic violations (Bergmann et al., 2015; but see Kasparian, 2015 for arguments against this claim). Despite the observation that L1 attrition of

morpho-syntactic phenomena is rare in speakers who acquired their L2 in adulthood, such cases have been reported (e.g., Sorace, 1993; Gürel, 2002; Iverson, 2012).

However, we still know quite little about what properties can undergo attrition and under what linguistic and extralinguistic conditions structural changes to L1 grammars can occur. To address this question as well as the rarity of L1 attrition, Hicks and Domínguez (2020) recently proposed the Attrition via Acquisition (AvA) model, a formal model of grammatical attrition which presents testable predictions for how changes to the L1 grammar may occur and uses principles of generative grammar in combination with psycholinguistic approaches on language acquisition to account for the rare occurrence of grammatical attrition.

The current study tests the predictions of the AvA on potential changes to L1 grammars in first generation immigrants who learned the L2 as adults and are living in the L2 environment. We focus on a phenomenon at the syntax-discourse interface, specifically the use of Clitic Left Dislocation (CLLD) in Romanian and its differences from Italian CLLD. To date, most previous studies on L1 attrition of discourse-syntactic phenomena have focused on the interpretation of anaphoric forms, like null and overt pronouns in null subject languages (Gürel, 2002; Tsimpli et al., 2004; Gürel and Yilmaz, 2011; Domínguez, 2013; Kaltsa et al., 2015; Chamorro et al., 2016; Miličević and Kraš, 2017) or pronominal and demonstrative pronouns in German (Wilson, 2009; Wilson et al., 2009), but attrition in CLLD has not previously been tested. To disentangle effects of L2 transfer on attrition from general L1 disuse we compare two groups of speakers with different L2s: one group whose L2 has CLLD (Italian), but uses this construction in different discourse contexts than Romanian, and a group of speakers whose L2 does not have CLLD (English). According to the AvA, grammatical attrition is only expected for properties where analogous forms exist in the L1 and the L2 but where these forms differ in their behavior due to differences in their feature specifications, as is the case for Romanians who acquired Italian as an L2. We furthermore examine

(2011). The IH predicts more optionality and variability in the performance of attriters compared to non-attriters for structures that require the integration of linguistic and non-linguistic information, like discourse information (Tsimpli et al., 2004; Sorace, 2011). Differences between groups are particularly predicted to be found in real-time processing, arguably due to reductions in working memory and processing efficiency (Rothman and Slabakova, 2011). Note that the IH is a theory of processing and assumes that attriters who migrated in adulthood do not have a grammar that qualitatively differs from monolingual non-attriters who speak the same L1. Instead, the grammatical errors that L1 attriters make are argued to be due to a momentary conflict between their two linguistic systems, causing instances of disfluency and optionality in the use of morpho-syntactic properties. Processing approaches to attrition more generally have argued that a lower frequency of activation can cause processing delays in bilinguals independently of L1-L2 differences (Gollan et al., 2005) or that it is cross-linguistic transfer in the form of competition and spread of activation from the L2 or other languages that can lead to less efficient processing (Marian and Spivey, 2003; Blumenfeld and Marian, 2007).

To examine whether attrition is due to a momentary conflict between the two grammars or a different grammatical representation, Chamorro et al. (2016) compared three groups of native Spanish speakers, two of which had been living in the United Kingdom for at least 5 years, and a group of Spanish native controls who had recently moved to the United Kingdom with very little knowledge of English. The two experimental groups differed in that the speakers in one group were recently re-exposed to Spanish only. Participants were tested on anaphora resolution of null and overt pronouns in Spanish using sentences like (1) where null pronouns (*pro*) have been shown to prefer subject antecedents while overt pronouns favor object antecedents. Two tasks were used, an online eye-tracking task and an untimed naturalness judgment task, as the authors assumed that online tasks measure real-time language processing and untimed offline tasks reflect knowledge representation.

1. La madre saludó a la chica cuando ella/*pro* cruzaba una calle con mucho tráfico.
The mother greeted to the girl when she/*pro* crossed a street with much traffic.
'The mother greeted the girl when she crossed a street with lots of traffic.'
- Adapted from Chamorro et al. (2016, Ex. 8).

the effects of language-external factors, such as age of immigration (including participants who immigrated during and after adolescence), relative L1 and L2 use, and L2 proficiency as factors contributing to attrition. Results from the current study are furthermore compared to those of an earlier study on L2 acquisition reported in Smeets (2023), as we observed interesting patterns of crosslinguistic influence in L2 acquisition and in L1 attrition that are alike.

2 Reduced processing efficiency or changes in grammatical representation?

Most previous attrition studies investigating linguistic phenomena involving the integration of discourse information into the syntax have been conducted in light of the Interface Hypothesis (IH; Sorace,

The results revealed that the monolingual and the re-exposed groups had faster go-past times in the critical region (pronoun or *pro*) when the overt pronoun had an object antecedent and when the null pronoun had a subject antecedent. The attrited group was faster when the pronoun matched the object rather than the subject, regardless of whether the pronoun was null or overt, suggesting a lack of sensitivity to pronoun type. No differences across groups were found in the offline naturalness rating task. The authors argue that the finding that the re-exposed group, who had been in a Spanish-only environment for only a week, did not differ from the Spanish monolingual controls suggests that it is unlikely that any permanent changes had occurred to the native grammars of these speakers. The authors furthermore take the absence of evidence for attrition in the offline task as evidence for the idea that attrition affects the ability to process interface structures, but does not affect knowledge representation. Crucially, however, earlier studies on the same linguistic phenomenon, specifically Tsimpli et al. (2004) on another pro-drop language (Italian), did find attrition

in the form of overgeneralization of overt pronouns in contexts where Italian monolinguals would use a null pronoun (i.e., with subject antecedents) using an offline antecedent selection task. Note, however, that it is quite difficult to know whether attrition affects grammatical representation on the basis of comparing performance in online versus offline tasks, as neither allows access to the brains of speakers to measure linguistic competence. As argued by White (2023), essentially all experimental tasks are measures of performance and “both offline and online measures can be used to determine the nature of linguistic representations, as well as processing considerations” (White, 2023, p. 334). We return to this discussion in Section 8 and show how the findings of the current study indicate differences in grammatical representation in the mental grammars of our attrited participants.

3 The importance of L1-L2 overlap for attrition

A well-supported finding in attrition research is that L1 forms that have no analogous forms in the L2 are more easily preserved than L1 forms that are in competition with L2 forms (Altenberg, 1991; Köpke, 1999, 2002; Pelc, 2001; Gürel, 2004, 2007; Paradis, 2007; Tsimplici, 2007). We can illustrate the importance of L1-L2 overlap using the interpretation of pronouns in Turkish and English as examined in Gürel (2002, 2004) and Gürel and Yilmaz (2011). Languages differ in the syntactic-semantic constraints on the interpretation of pronouns. The Turkish pronoun *o* functions differently from English pronouns *him/her/they*: while English allows bound interpretations (*he* can refer to *Burak* in (2)), this reading is not possible for the Turkish pronoun *o*.

2. Burak_i [o_{i/k}-nun zeki ol-duğ-u]-nu düşün-üyor.
 Burak s/he-GEN intelligent be-NOM-3SGPOSS-ACC think-PRG
 ‘Burak_i thinks that s/he_{i/k} is smart’

In addition to overt pronouns, Turkish also allows null pronouns and the anaphoric pronominal *kendisi* in subject position, while no such counterpart exists for English. To examine the potential effect of English on the interpretation of Turkish *o*, Gürel (2002, 2004) tested Turkish native speakers who were near-native speakers of English and immigrants to North America, as well as native controls in Turkey. The author reports that while the L2 English group and the Turkish monolingual control group did not behave any differently in their interpretation of null pronouns and *kendisi*, the L2 English group chose a bound interpretation of *o* significantly more often than the control group. Gürel (2002, 2004) therefore concludes that competition with an L2 form is needed for attrition to occur. Following the same reasoning, the findings in Tsimplici et al. (2004) also support the idea that structural overlap is needed for attrition to occur. In pro-drop languages like Italian (and Spanish, discussed in Section 2), null pronouns (*pro*) refer to subject antecedents, typically the topic of the sentence, while overt pronouns tend to refer to object antecedents. Following the syntactic analyses in Cardinaletti and Starke (1994) and Grimshaw and Samek-Lodovici (1998), Italian pronouns are argued to be specified for [+Topic shift]. Overt pronouns in non pro-drop languages like English can refer to subjects and objects

and are not specified for this discourse feature. Because English and Italian both use overt pronouns, L2 English can cause Italian overt pronouns to become optionally underspecified for [+Topic Shift] in the grammars of Italian native speakers who have become dominant speakers of English. In consequence, attriters may over-accept and use overt pronouns in their pro-drop L1, allowing them for both subject and object antecedents. This is exactly what Tsimplici et al. (2004) found: attriters overgeneralized overt pronouns in contexts where Italian monolinguals would use a null pronoun. The interpretation of Italian null pronouns, however, was not affected, as there is no L2 counterpart.

Similarly, Chamorro et al. (2016) tested potential attrition on the use of the object marker *a*, which in Spanish is required to precede a direct object when the object is animate and specific. The results show no signs of attrition in the Spanish of near-native speakers of English in the United Kingdom. Because the participants were the same as in Chamorro et al. (2016) on the interpretation of over pronouns, Chamorro and Sorace (2019) compare the performance in the two studies and attribute the different results to the fact that the distribution of null and overt pronouns involves the external interfaces while the use of the object marker *a* is driven by semantic factors (animacy and specificity) and therefore involves the internal interfaces. However, an alternative explanation for the lack of attrition with the use of *a* is the absence of an L2 analogous form. Crucially, English does not allow differential object marking and therefore L2 properties cannot possibly influence the [+animate] and [+specific] feature of the L1 grammar. English does have overt pronouns but their use differs from pronouns

in null-subject languages, causing competition between L1 and L2 forms.

The need for L1-L2 analogous forms as a prerequisite for attrition has also been supported by Iverson (2012). Using online and offline acceptability and interpretation judgment tasks, Iverson (2012) examined the grammar of a native Brazilian Portuguese speaker who by the time of testing almost exclusively spoke his L2 Chilean Spanish. The participant was tested on a range of phenomena at the external interfaces, internal interfaces and phenomena pertaining to narrow syntax. Iverson (2012) found that the main predictor for attrition was not whether the property pertained to the external interfaces, as would be predicted by the IH, but whether Brazilian Portuguese and Chilean Spanish share properties. The author furthermore argues that the speaker's grammar is qualitatively different from monolingual L1 grammars, as his grammar reflected convergence with the L2 grammar in all linguistic structures where the L1 and the L2 differed.

To conclude, research has focused on whether attrition can cause structural changes to native grammars and if so under which conditions. Although grammatical attrition is likely to be rare, syntactic restructuring has been attested and is more likely to take place as the result of long-term co-activation of a

language system that has analogous forms. More research with a broader variety of linguistic structures, language combinations and a combination of various experimental tasks is needed to provide additional insights. The current study tests whether the

couch” does not have an antecedent (it is not mentioned in the immediate discourse).¹ While in English the dislocated object is not doubled by a preverbal clitic in either (3) or (4), in Romanian and in Italian such sentences can be expressed using

3. Topic Fronting

Q: What did you do with the couch and the table?

A: [The couch]_i I put *t_i* in the living room, but the table broke during transportation.

4. Focus Fronting

Q: You put the table in the living room, right?

A: [THE COUCH]_i I put *t_i* in the living room, not the table.

existence of analogous forms in the L1 and L2 can lead to grammatical restructuring of the L1 using another phenomenon at the syntax-discourse interface, namely the use of Clitic Left Dislocation. We are comparing the L1 Romanian grammars of speakers whose L2 does have an analogous form (Italian) to speakers whose L2 does not use CLLD (English). The properties of CLLD in Romanian and Italian will be discussed in the next section.

4 Clitic left dislocation in Romanian and Italian

This paper focuses on two types of A-bar movement of an object into the left-periphery: Contrastive Topic Fronting and

Clitic Left Dislocation. This section examines the cross-linguistic differences associated with Clitic Left Dislocation (CLLD), an example of which is shown in (5a) for Romanian and in (5b) for Italian.

In both languages, the dislocated object *the couch* is doubled by a preverbal clitic. However, different conditions underlie the use of clitics with left dislocation in Italian and Romanian, as different features are involved. The two relevant features are *specificity* ([± specific]) and *discourse anaphoricity* ([± anaphor]), the exact mechanisms for which are discussed in Smeets (2022, 2023). In Romanian, only dislocated objects that have a specific referent participate in CLLD [Cornilescu and Dobrovie-Sorin (2008); Avram and Coene (1999); Smeets (2022, 2023)]. In (5), the speaker has a specific couch in mind. However, if we look at clitic use in a scenario where the dislocate

5. [+specific, +anaphor]

Q: What did you do with the couch and the table?

a. [Canapeaua]_i am pus-*(o) în sufragerie, dar masa s-a rupt în timpul transportului.
Couch.DEF have put-CL.ACC.F.SG in living-room but table.DEF REFL-IS broken in time transportation.
‘The couch I put in the living room, but the table broke during transportation.’

b. [Il divano]_i *(l’) ho messo in soggiorno, ma il tavolo si è rotto durante il trasporto.
The couch CL.ACC.M.SG have put in living-room but the table REFL is broken during the transportation
‘The couch I put in the living room, but the table broke during transportation.’

Contrastive Focus Fronting, as shown in (3) and (4) respectively.

In both sentences, the fronted object receives a contrastive interpretation (López, 2009) where “the couch” is contrasted to “the table” mentioned in the previous sentence. Following López, the two constructions in (3) and (4) can be differentiated by the discourse property [± anaphor]. In (3), the object is an example of a discourse anaphor, as the dislocate “the couch” has an antecedent in the immediate discourse (a local antecedent, see Villalba, 2000) and the answer elaborates on the previous sentence by contributing new information about what happened to the couch. In (4), the dislocate “the

is non-specific, as is the case for *a red skirt* in (6) and (8), we see that clitics are not allowed in Romanian. Italian CLLD, on the other hand, is used with both specific and non-specific objects (compare (5b) to (6b)), as Italian CLLD is not constrained by specificity.

¹ Focus fronting is used most commonly to express corrective focus whereby the object is placed in a position of prominence with the goal to provide a correction.

6. [−specific, +anaphor]
Q: Did you find a red skirt and a pair of boots?
a. O fustă roșie (*o) caut deja de două luni, dar am găsit o pereche de ghete negre.
 a skirt red CL.ACC.F.3SG search already for two months but have.1SG found a pair of boots black
 ‘I’ve been looking for a red skirt for two months, but I did find a pair of black boots.’
b. Una gonna rossa *(la) cerco già da due mesi, però ho trovato un paio di stivali neri.
 a skirt red CL.ACC.F.3SG search.1SG already since two months but have.1SG found a pair of boots black.
 ‘I’ve been looking for a red skirt for two months, but I did find a pair of black boots.’

In Italian, however, CLLD is restricted to discourse topics and cannot be used with contrastive focus fronting, as the ungrammaticality of the clitic shows in (7b). Romanian CLLD, on the other hand, is used with both topic ([+anaphor]) and focus ([−anaphor]) fronting, as shown in (5a) and (7a).

or nouns in Romanian, Italian and English. In Romanian the distribution of clitics is dependent on whether the fronted object is [± specific] and whether an (indefinite) noun is interpreted as specific or non-specific depends on whether there is a specific referent available in the discourse context. Similarly, whether or

7. [+specific, −anaphor]
Q: You put the table in the living room, right?
a. CANAPEAUA am pus-*(o) în sufragerie, nu masa. Masa s-a rupt în timpul transportului.
 couch-DEF have.1SG put-CL.ACC.F.3SG in living-room not table-DEF table-DEF REFL-has broken in time transportation
 ‘The couch I put in the living room, not the table. The table broke during the transportation.’
b. Il DIVANO (*l) ho messo in soggiorno, non il tavolo. Il tavolo si è rotto durante il trasporto.
 The couch CL.ACC.M.SG have.1SG put in living-room not the table. the table REFL is broken during the transportation.
 ‘The couch I put in the living room, not the table. The table broke during the transportation.’

In the context in (8) neither Italian nor Romanian uses a clitic. A clitic is not allowed in Italian because the fronted object is [−anaphor] and it is not allowed in Romanian because the dislocate is [−specific].

not a constituent is discourse anaphoric requires the reader or listener to keep information from the previous discourse in working memory. Therefore, in both Romanian and Italian, the presence or absence of a clitic depends on changing contextual

8. [−specific, −anaphor]
Q: Weren’t you looking for a red sweater? I saw some nice ones at H&M.
a. O FUSTĂ roșie o caut, nu o cămașă roșie.
 A SKIRT red CL.ACC.F.3SG seek-for.1SG not a sweater red.
 ‘I am looking for a red skirt, not a red sweater.’
b. Una GONNA rossa *(la) cerco, non una maglietta rossa.
 A skirt red CL.ACC.F.3SG look-for.1SG, not a sweater red.
 ‘I am looking for a red skirt, not a red sweater.’ Smeets (2023), examples 4–7.

Although specificity is typically assumed to be a semantic feature on noun phrases, specificity is not marked on determiners

information. The distribution of clitics in Romanian and Italian is summarized in Table 1.

TABLE 1 Distribution of resumptive clitics in Italian and Romanian.

		[+ anaphor]		[− anaphor]		Property
		[+ specific]	[− specific]	[+ specific]	[− specific]	
1	Italian	✓	✓	χ	χ	anaphoricity
2	Romanian	✓	χ	✓	χ	specificity

TABLE 2 Expected use of CLLD per condition according to the predictions of the AvA.

		[+ anaphor]		[– anaphor]	
		[+ specific]	[– specific]	[+ specific]	[– specific]
1	Romanians in Italy	✓	✓	✓	✗
2	Romanians in anglophone countries	✓	✗	✓	✗

5 The attrition via acquisition model: predictions for CLLD

The current study examines the use of CLLD by Romanian first generation immigrants to either Italy or an anglophone country. The hypotheses and results will be interpreted in light of the Attrition via Acquisition (AvA) model (Hicks and Domínguez, 2020), a formal model of grammatical attrition which provides a testable hypothesis for the conditions where L2 properties may change mature L1 grammars.

The model is developed within the generative framework, which assumes that differences between languages are expressed in the features they select from an innately available universal set of features and the way they apply those features to lexical items and morphemes (as in the Minimalist Program, Chomsky, 2000, 2001 et seq.). The AvA addresses the question of whether formal features that are set in early childhood can be reset due to influence of L2 features in speakers with reduced access to and use of their L1. In order to explain changes to feature representations, Hicks and Domínguez (2020) elaborate on the Feature Reassembly Hypothesis (FRH, Lardiere, 2008), a prominent theory in generative language acquisition research which examines the role of L1 transfer into L2 grammars to explain relative difficulty and success in L2 acquisition. Specifically, the FRH is developed around the fine-grained differences across languages on how they encode grammatical features. The FRH predicts that L2ers at the initial stage transfer the features associated with specific lexical items into the L2 grammar. The task of an L2er then involves adjustments to features on morphemes or lexical items that were incorrectly transferred from the L1 grammar. Applying the FRH to attrition, the Attrition via Acquisition (AvA) model argues that grammatical attrition consists of adjustments to L1 features on individual morphemes that are transferred from the L2 in situations of L1-L2 overlap. The AvA therefore predicts L1 grammatical attrition to be possible when there are analogous forms in the L1 and L2 that yet differ to some extent, in line with the findings discussed in Section 3.

Crucially, however, while the presence of L1 features in L2 grammars is extremely common and well-attested, the presence of L2 features in L1 grammars is certainly much rarer. To explain the rarity of grammatical attrition, Hicks and Domínguez follow Lidz and Gagliardi (2015)’s theory on L1 acquisition to discuss how grammatical properties of a grammar that has reached maturity can become less stable and open for the intake of new grammatical properties. The AvA model assumes a unified mechanism for acquisition and attrition where L1 attrition engages in the same acquisitional mechanisms as L1 and L2 acquisition. The theory decouples linguistic input from acquisitional intake, which is the information the mind actually extracts from the input. Hicks and Domínguez propose that when extensive L2 input is accompanied with a reduction in L1 input, the so-called ‘inference engine’ may be reopened to take in new features that update the existing L1 grammar. L2 interference, a prerequisite for the eventual alternation of L1 representations, can only occur for linguistic phenomena where there is overlap between the L1 and the L2 but where the L2 assumes different

values (features) for corresponding linguistic items. Specifically, Hicks and Domínguez (2020, p. 156) predict grammatical attrition to be possible under the following circumstances:

- The L2 is close, yet not identical, to the speaker’s first language. Specifically, the L1 and the L2 allow a certain syntactic construction but use them in different situations. Hence, prolonged exposure to the L2 can alter L1 feature-form mappings, but only if the L2 allows for the same syntactic construction as the L1.
- The changes in the L1 grammar do not involve a loss of options from the L1 grammar, or replacement of L1 features by L2 features. Instead, options from the L2 grammar supplement the existing grammar, leading to L1 restructuring.

As pointed out by Schmid and Köpke (2017b), a feature reassembly approach to L1 loss where L2 features influence the feature bundles of the L1 has previously only been applied to contexts of heritage language acquisition, where developing L1 grammars whose feature representations are shown to be weaker or incomplete are affected by L2 features (Putnam and Sánchez, 2013) and to contexts of contact-induced change involving two minimally different varieties of the same language (Cuban and Peninsular Spanish, see Domínguez and Hicks, 2016). While the AvA predicts that grammatical attrition is favored when the L1 and L2 are typologically more similar, it also predicts that similar (but not identical) comparative behavior in the L1 and L2 provides a sufficient condition for attrition to occur. Hicks and Domínguez (2020) are the first to apply the Feature Reassembly approach to attrition in late L2 learners of a different (mutually unintelligible) language.

Specifically, the authors illustrate the applicability of the AvA model with data from previously reported findings on the realization and interpretation of pronominals in adult first generation immigrants. As discussed in Section 3, overt pronouns in pro-drop languages like Italian and Spanish have been argued to have a [+Topic shift] feature, where the use of an overt pronouns indicates an interpretation away from the discourse topic (typically the subject). Overt pronouns in non pro-drop languages like English do not have a discourse feature, as pronouns are used with both topic and non-topic antecedents. The existence of overt pronouns in both languages, albeit used in different contexts, meets the requirement of the AvA for grammatical attrition to be possible. The properties of the overt pronoun of the L2 can be transferred onto the L1 and affect the use and interpretation of pronominals in attrited native speakers of a pro-drop language. The presence of overlap in the use of overt subject pronominals in both English and Italian/Spanish can cause attriters to associate the feature specifications of English pronominals (which lacks a [+Topic Shift] feature) with the corresponding pronominal of Italian/Spanish. In consequence, overt pronouns are also used in contexts where there is no topic shift. The authors argue that “continued processing of L2 input that invokes both UG and the L1 in updating the advanced L2 grammar allows for the possibility that acquired morphosyntactic features of the relevant L2 lexical item ‘update’ the L1 grammar” (Hicks and Domínguez, 2020, p. 157).

TABLE 3 Background information, showing mean (range) for age, age of arrival, length of residence, relative language proficiency and language use.

	Italian monolinguals	Romanian monolinguals	L2 English	L2 Italian
Number of participants	18	17	23	37
Age	33.4 (24–53)	33.4 (24–51)	51 (22–69)	39.44 (23–58)
Age of arrival			29.7 (18–47)	23.03 (11–40)
Length of residence			21.3 (4–21)	16.4 (3–20)
Relative language proficiency			3.65	3.53
External language use			3.16	2.82
Internal language use			3.43	3.38

To examine the validity of the AvA and to further improve the model, it is important to apply the predictions to new linguistic contexts and language combinations. To date, the application of the AvA has mostly focused on studies examining the interpretation of overt pronouns in L1 pro-drop languages [but see recent work by [Baker \(2024\)](#) who examined a variety of morpho-syntactic phenomena, finding different types of changes to L1 grammars for some speakers and for some linguistic phenomena]. It is relevant to note that for the use of overt pronouns in pro-drop languages, it is hard to convincingly conclude that the attested attrition effects are due to L2 transfer in the absence of another group whose L1 and L2 work the same. As pointed out by [Montrul \(2008\)](#), the over-acceptance of overt pronouns with topic antecedents can also be due to the alleged complexity associated with the syntax-discourse interface, as predicted by the IH. Furthermore, attrition effects independent of L2 transfer have also been attested in the form of simplification, where marked forms are replaced by unmarked forms ([Schmid, 2002](#), p. 13). Applying this reasoning to the use of subject pronouns, the overuse of overt pronouns in contexts without topic shift could also be the result of those speakers resorting to unmarked values by removing the [+topic shift] feature from overt pronouns. For this reason, it is important to also include an experimental group where L2 transfer cannot occur, which may be impossible for the interpretation of overt pronouns, assuming overt pronouns exist in all languages. The cross-linguistic differences between Romanian and Italian and the lack of CLLD in English form an ideal test case to disentangle L2 transfer effects as predicted by the AvA model from other factors potentially causing changes to attrited grammars.

6 Current study

Our discussion so far has focused on the effects of L2-driven factors on attrition and the importance of analogous forms in the L1 and the L2 for attrition to occur. To further test the AvA, we compare the use of CLLD by Romanian native speakers who are living in an English-dominant environment (no L2 transfer possible) to Romanians living in Italy (L2 transfer possible), as attrition is predicted to be possible only for the latter group. Specifically, L2 options are predicted to supplement the L1 grammar. This means that grammatical attrition does not involve a complete loss of L1 forms but a fluctuation between the grammatical options from the L1 and the L2.² For the use of CLLD, attriters who are

L2 speakers of Italian are expected to supplement their L1 Romanian grammar with options available in Italian and therefore allow clitics when the fronted object is a non-specific topic. Although the AvA would not predict any changes to the mental representation of Romanians in anglophone countries, data from the L2 English group can provide insights into whether there are causes to attrition that are independent from L2 transfer, for example due to reduced activation of the L1 grammar possibly which may cause inconsistent or inefficient processing. The predictions following the AvA are summarized in [Table 2](#).

Since we are interested in examining whether attrition affects knowledge representation, we tried to reduce processing difficulties as much as possible by using two untimed tasks, an Acceptability Judgment and a Written Elicitation task. We expect Romanians in Italy to accept clitics with non-specific topics in the Acceptability Judgment task and to use clitics in this context in the Written Elicitation task. We furthermore expect them to continue to use and accept clitics with specific foci.

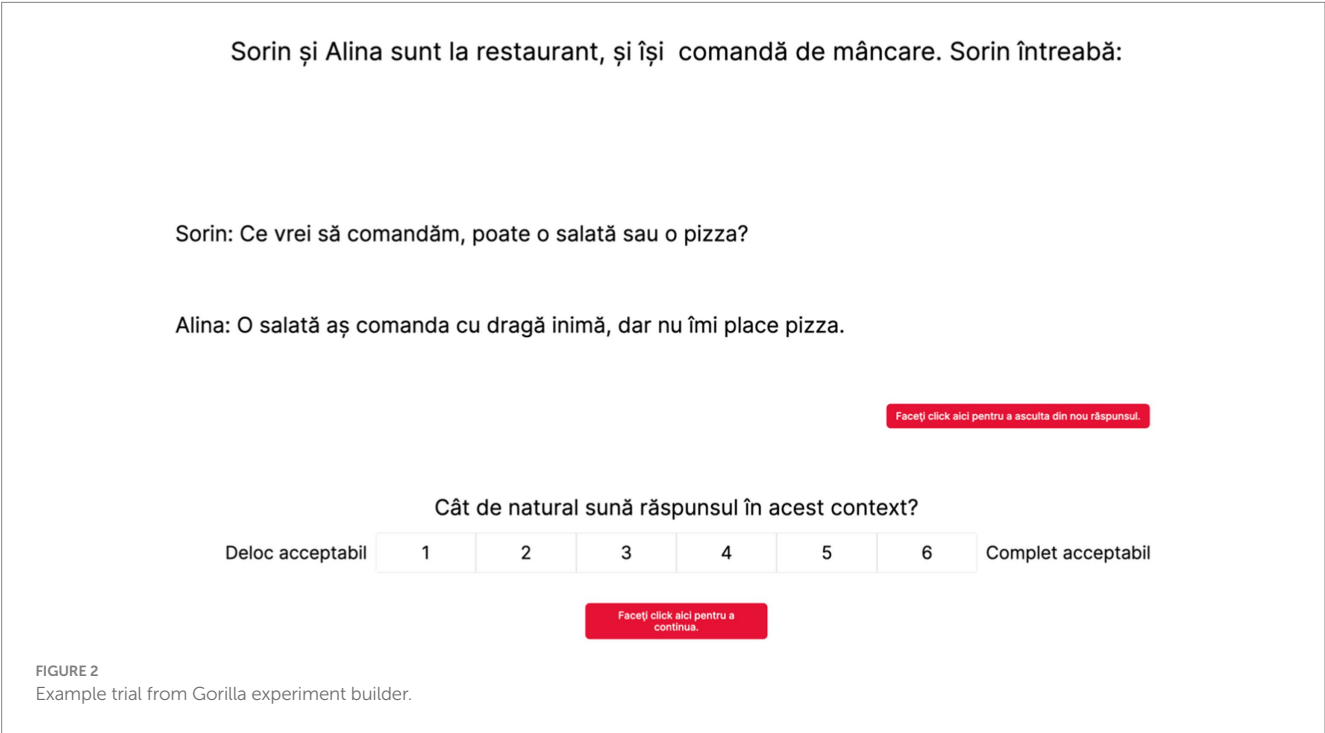
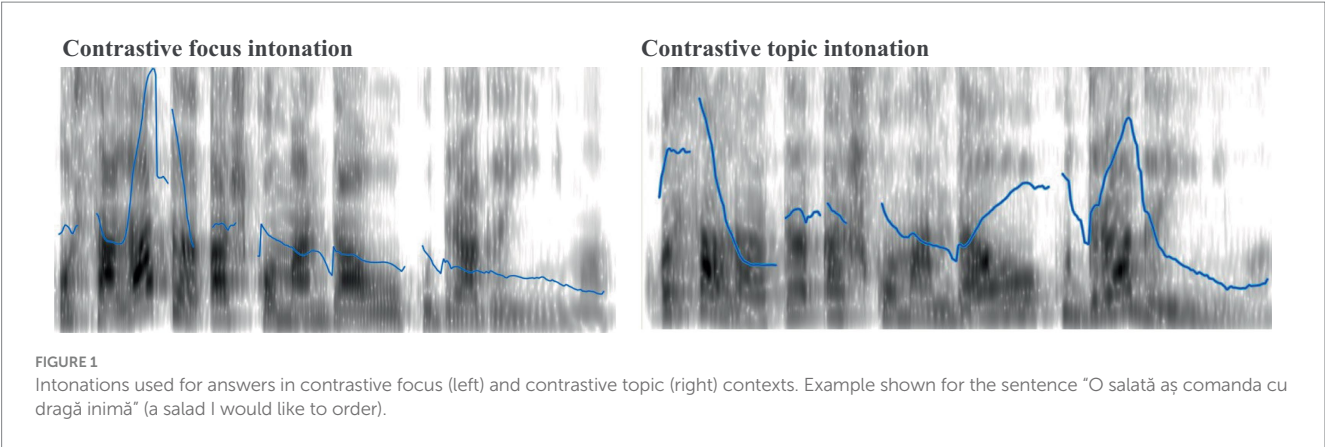
6.1 Methods

6.1.1 Participants

A total of 95 participants completed the experimental tasks, either in Italian or in Romanian (see participant details and background information in [Table 3](#)). Participants were recruited by research assistants who were in-group members of Romanian immigrant communities in Italy or the US and Canada. Prior to participation it was ensured that none of the participants spoke another language with CLLD (e.g., Spanish, Greek or Bulgarian). In addition to the Acceptability Judgment task and the Written Elicitation task, which were always presented in this order, participants filled out an extensive background questionnaire adapted from [Schmid and Dusseldorp \(2010\)](#) to gain more information on extralinguistic factors that may affect attrition rate. Note that the questionnaire included a question asking about participants' current age, the age at which they started to acquire the L2 and their length of residence in the country where the L2 is spoken. Since English is a language that is widely taught in elementary schools as a second language, the age of onset of the L2 English group was much lower than that of the L2 Italian group. However, since any signs of L1 attrition likely only start to occur when immersed in an environment where the L2 is spoken, we used the age of arrival instead, by subtracting the length of residence from their current age. The questionnaire furthermore

² Interestingly, Pablo, the native Portuguese speaker in Chile studied in [Iverson \(2012\)](#) did show a complete loss of L1 forms. His case is, however, quite unique,

as the speaker completely stopped using his L1.



consisted of several questions related to language use as well as those related to the participant’s language and cultural attitudes. Following the suggestions in Schmid (2011, p. 214), it is desirable to reduce the number of responses from each individual question by taking an average over various questions that can be grouped together. Note, however, that there are no specific guidelines available for coding questions into such broader categories. I decided to group together various questions that relate to relative language proficiency, questions that ask about external language use both inside and outside the home and a third group of questions that relate to internal language use, which is argued to be indicative of someone’s linguistic and cultural affiliation (Schmid, 2011). Specifically, *Relative language proficiency* indicates an average score participants gave to nine questions, listed below, related to self-reported relative language proficiency. Participants were asked to indicate whether they were more proficient in the L1 (Romanian) or the L2 (Italian/English) using a 1–5 scale where 1 indicates “only Italian/English,” 2 “mostly Italian/English,” 3 “both,” 4 “mostly Romanian” and 5 “only Romanian.” An average score of 3.0 indicates that speakers had equal proficiency in both languages.

1. In which language do you have the largest vocabulary?
2. In which language do you have no pronunciation issues?
3. Which language are you able to use/understand intuitively?
4. In which language are you familiar with various dialects/slang?
5. In which language do you have an intuitive feeling of what is “correct”/ “incorrect”?
6. Into which language are you able to translate?
7. In which language can you understand and make jokes?
8. In what language do others consider you a native speaker?
9. In which language can you express yourself more easily?

External language use shows the average response to seven questions that asked participants what language(s) they speak with (1) friends, (2) daily basis, (3) partner, (4) pets, (5) work, (6) relatives, and (7) at the store. The score for *Internal language use* consists of six questions asking participants in which language they (1) think, (2) dream, (3) count, (4) swear, (5) use when emotional and (6) use when upset. These questions used the same five-point rating Likert scale. The average response to each group of questions and each L2 group is shown in Table 3.

As can be observed in Table 3, the length of residence in the L2 country is somewhat longer for the L2 English group, although speakers in both groups, on average, have been living in the country where the L2 is spoken for a substantial number of years. Note that the L2 Italian group included speakers who immigrated during early adolescence, while the earliest age of arrival for the L2 English group is 18. We will get back to effects of onset of immigration in Section 7. The two groups are highly comparable in their responses to the sociolinguistic variables, with both groups stating to only be slightly more proficient in their L1 than their L2 and to use their native language slightly more internally (which has been argued to reflect their cultural and linguistic identity).

6.1.2 Experimental tasks

In the Acceptability Judgment task, participants were asked to judge the acceptability of sentences with left dislocated word orders on a 6-point rating scale where 1 indicated highly unacceptable and 6 highly acceptable in the context in which they were presented. Experimental trials were like the examples in (5) to (8) in which the contexts, questions and answers were presented in Italian for the Italian monolingual control group and in Romanian for the Romanian monolinguals and the two attriter groups. All experimental trials can be found.³ The experiment consisted of 64 target items which differed by three factors: *Discourse* and *Clitic*, which are within-item factors, and *Specificity*, which is a between-item factor. The factor *Discourse* had two levels: in the Topic context, as in (5) and (6), the left dislocated object was discourse anaphoric and in the Focus context, as in (7) and (8), the fronted object was not discourse anaphoric. For *Specificity*, the fronted object was either specific, as in (5) and (7), or non-specific, as in (6) and (8). For *Clitic*, there was either a clitic or no clitic.

Stimuli were presented both in written and auditory form to ensure that participants processed the sentences with the intended intonation. To ensure the pronunciation of the question-and-answer pairs was most natural given the context provided, each full experimental trial, including context, question and answer, was recorded by both a female

and a male native speaker of Romanian. To create the experimental trials, recordings of the contexts and the answers were then taken from the female voice and the questions were taken from the male voice (in alignment with the gender of the characters in the story). The female was also a linguist who helped ensure that the pronunciation used was most natural given the intended interpretation. As illustrated with an example in Figure 1, fronted constituents in the Focus contexts received a high tone (H*) followed by a default low tone [following Jitcă et al. (2015) who identified this intonational pattern for Romanian contrastively focused constituents]. Contrastive topic configurations were associated with a ‘rise-fall-rise’ intonation, as argued by Buring (2002) for English and in agreement with judgments from native Romanian research assistants for Romanian.

Each experimental trial was presented as follows: the context automatically appeared, after which the question-and-answer pair was shown, both in written and spoken form. Thereafter, the acceptability judgment scale appeared asking participants to rate on a scale of 1 to 6 how natural the answer sounded in the context provided. An example is shown in Figure 2.

In the Written Elicitation task (WET), participants were asked to complete sentences that were partially left blank and were instructed to complete the sentence using a part of a word or alternatively one, two or three words. The parts left blank aimed at eliciting a verb alone or a clitic and a verb. The WET consisted of 16 target items, four of each in the four possible conditions as illustrated in (9)–(12), varying the factors [\pm specific] and [\pm anaphor]. For each example in (9)–(12), the first answer illustrates a target sentence for Romanian and the second for Italian. In the actual experiment, the context, question and answer were of course completely shown in either Romanian or Italian. The experimental items were randomly interspersed with 53 filler items, where participants were asked to complete the sentences using tense, number and person inflections, prepositions and determiners. Each item started with a short context followed by a question-answer pair.

9. [+ specific, +anaphor]

Livio is looking for someone who can take his granny's cat and dog as she cannot take care of them anymore. Livio asks Silvia:

Q: Would you maybe want to adopt the cat or the dog?

A: Pisica o voi adopta/ as adopta-o/ o pot adopta cu drag, dar nu avem loc pentru un câine.

A: Il gatto lo adotterei/lo prenderai volentieri, ma non abbiamo spazio per un cane.

The cat (CL) would/will/can.1sg adopt/take happily, but I do not have space for a dog.

10. [+ specific, –anaphor]

Anna and Beatrice are talking about Lea and Gianni who recently got married. Anna says to Beatrice:

Q: They have visited the Virgin Islands if I remember correctly.

A: Insulele MALDIVE le-au vizitat în luna de miere, nu Insulele Virgine.

A: Le MALDIVE hanno visitato per il viaggio di nozze, non le isole Vergini.

The Maldives (CL) have.3pl visited for the honeymoon, not the Virgin Islands.

11. [–specific, +anaphor]

Alessandra is in the library but she is not sure what she wants to read and she goes to the librarian to ask for recommendations. The librarian says:

Q: Would you like to read a book about airplanes or one about cars?

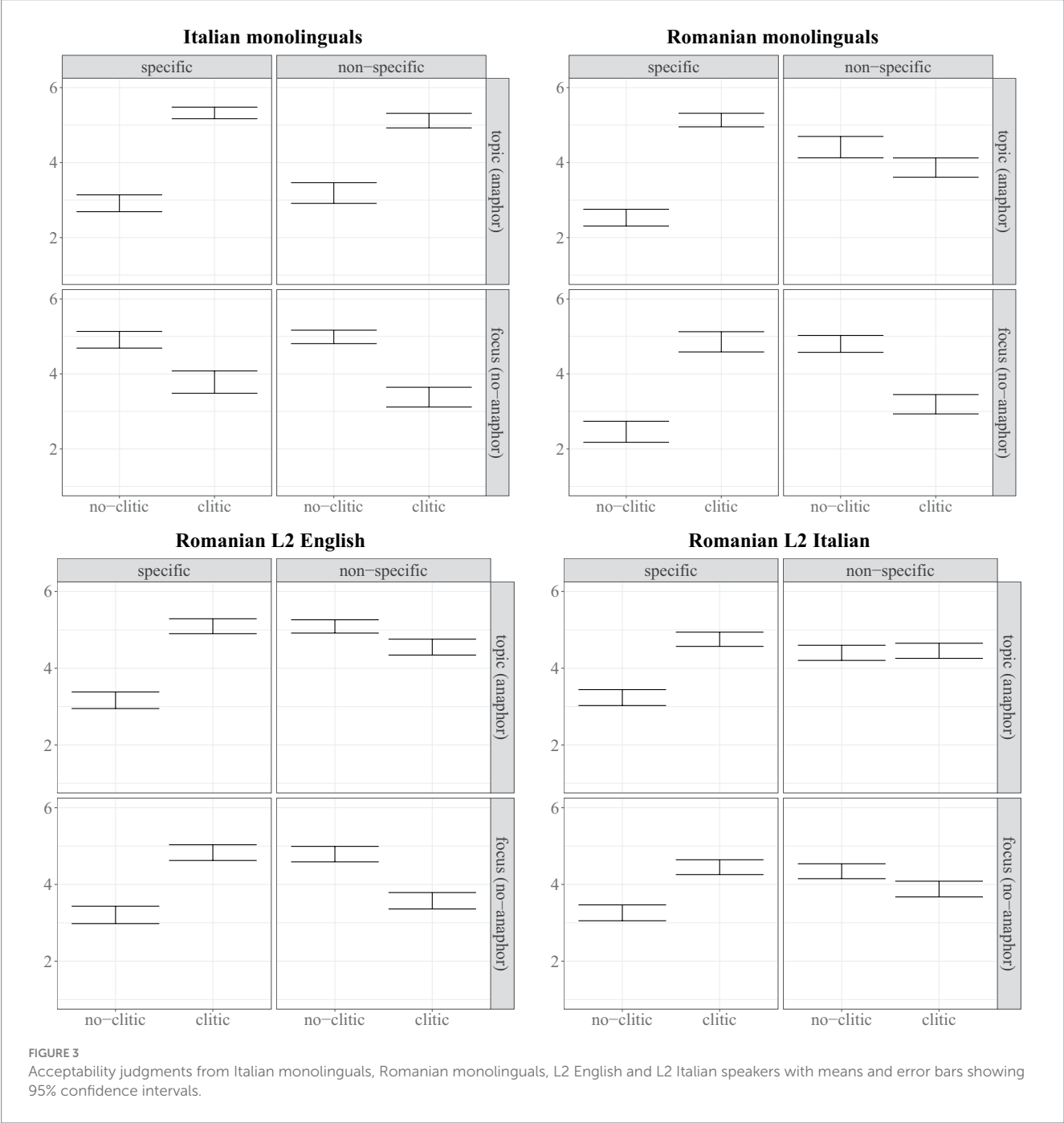
A: O carte despre avioane as citi cu plăcere, dar mașinile nu mă interesează.

A: Un libro sugli aeroplani lo leggerei con piacere, ma le macchine non mi interessano.

A book about airplanes (CL) will/would.1sg read with pleasure, but I am not interested in cars.

³ <https://osf.io/5xmfw>

12. [–specific, –anaphor]
Elena will go shopping this weekend because she has a date. Her friend tries to be helpful and says:
Q: Weren't you looking for a black shirt? I saw some cute ones at Zara.
A: O ROCHIE neagră caut, nu o cămașă neagră.
A: Un VESTITO nero cerco, non una maglietta nera.
A black dress (CL) search.1sg., not a black shirt.



6.2 Results

6.2.1 Acceptability judgment task

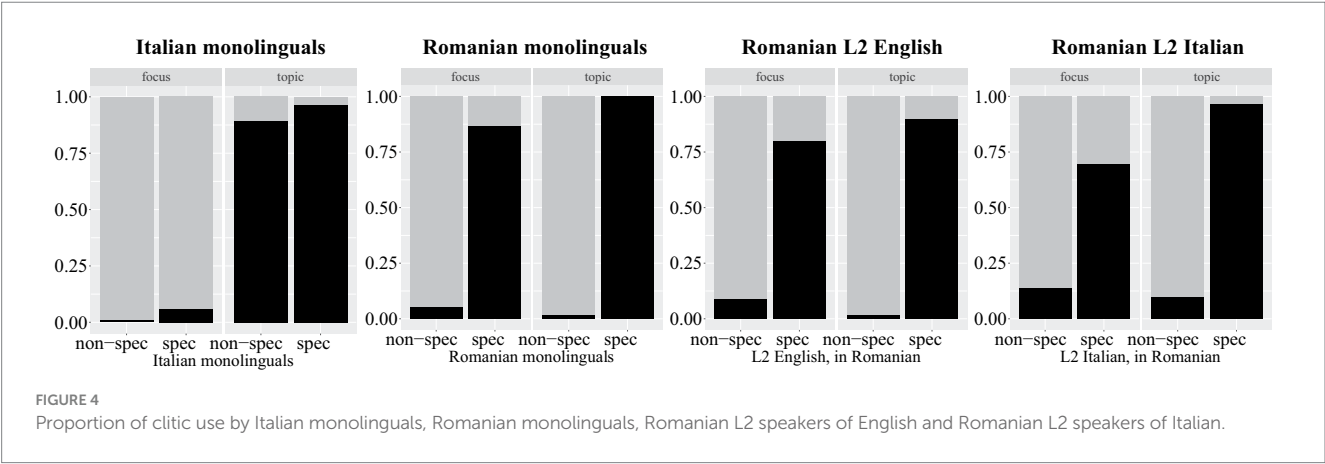
The results were plotted for each group separately to examine systematic patterns within the (interlanguage) grammars themselves

(following Bley-Vroman, 1983). All felicity ratings were analyzed using cumulative link mixed effects models (Christensen, 2014). The models include fixed effects for the categorical predictors *Clitic*, *Specificity*, *Discourse* and their interactions and random effects for Participant and Item. Each contrast was centered using sum coding.

TABLE 4 Acceptability judgments from Romanian monolinguals, L2 English and L2 Italian speakers.

Effects on acceptability judgment				
	Italian mon	Romanian mon	L2 English	L2 Italian
Predictor	Estimate(SE), <i>p</i> -value	Estimate(SE), <i>p</i> -value	Estimate(SE), <i>p</i> -value	Estimate(SE), <i>p</i> -value
Clitic No-Yes	0.67 (0.24), <i>p</i> = 0.006	1.33 (0.41), <i>p</i> = 0.001	3.20 (0.46), <i>p</i> < 0.001	2.27 (0.45), <i>p</i> < 0.001
Specific Spec-NonSpec	−0.10 (0.24)	0.89 (0.50)	2.92 (0.54), <i>p</i> < 0.001	1.61 (0.40), <i>p</i> < 0.001
Anaphor Top-Foc	−0.04 (0.19)	−0.49 (0.28)	0.01 (0.18)	0.05 (0.18)
Clitic No-Yes: Specific Spec-NonSpec	−0.92 (0.47), <i>p</i> = 0.048	−7.48 (1.36), <i>p</i> < 0.001	−4.15 (0.65), <i>p</i> < 0.001	−2.19 (0.45), <i>p</i> < 0.001
Clitic No-Yes: Anaphor Top-Foc	−6.63 (1.16), <i>p</i> < 0.001	−1.29 (0.51), <i>p</i> = 0.012	−0.55 (0.30)	−0.53 (0.27), <i>p</i> = 0.05
Specific Spec-NonSpec: Anaphor Top-Foc	−0.48 (0.36)	−0.23 (0.90)	−0.46 (0.32)	−0.09 (0.26)
Clitic No-Yes: Specific Spec-NonSpec: AnaphorT-F	−0.00 (0.56)	−0.03 (0.11)	−0.47 (0.47)	−0.25 (0.38)

Each column shows the estimate for each predictor with the standard error in parentheses. Significant effects are shown in bold, and *p*-values are provided.



Random-effect slopes were based on a maximal model that allowed convergence, following [Barr et al. \(2013\)](#).

[Figure 3](#) shows the results from the Italian and Romanian monolinguals, Romanians in anglophone countries, and Romanians in Italy, and [Table 4](#) the outcomes of the cumulative link mixed effect models for each of the four groups. Each line in the table starts with the label of the predictor in bold (e.g., Clitic) followed by the levels that were contrasted. For example Clitic No-Yes shows the effect of the absence vs. presence of the clitic, where the first level shows the baseline.

Discourse has the strongest effect in Italian, as shown by the interaction between Clitic and Discourse (Topic vs. Focus), and Specificity has the strongest effect in Romanian monolinguals, as shown by the interaction between Clitic and Specificity [see [Smeets \(2023\)](#)] for more details on the monolingual results, including a model that directly compares the two groups showing they are significantly different. The same significant interaction between Clitic and Specificity is found for the L2 English and L2 Italian groups, who rate clitic sentences as more acceptable with specific objects and no-clitic sentences as more acceptable with non-specific objects. Furthermore, there is a marginally significant interaction between Clitic and Discourse for the L2 Italian group, suggesting an effect of Discourse on the ratings of clitic vs. no-clitic sentences. As shown in [Figure 3](#), L2 Italian speakers as a group rate the acceptability of clitic and no-clitic sentences in Romanian as equally acceptable with fronted non-specific topics.

To summarize, Romanian monolinguals as well as L2 English and L2 Italian speakers rate clitic sentences as more acceptable than no-clitic sentences when the fronted object is specific. Additionally, Romanians in Italy rate clitic and no-clitic sentences as equally acceptable in the non-specific topic condition while the other native Romanian groups rate no-clitic sentences as more acceptable in this context. This result is suggestive of L2 influence as clitics with non-specific topics have become acceptable in the L1 of Romanians who acquired Italian. In Section 7 we elaborate on the L2 Italian findings by looking at differences across speakers in this group and show that the equal rating of clitic and no-clitic sentences with non-specific objects is driven by some speakers rating no-clitic sentences as more acceptable (in line the Romanian grammar) and others rating clitic sentences as more acceptable (in line with the Italian grammar).

6.2.2 Written elicitation task

All answers in the Written Elicitation task were manually coded by a native Romanian linguist, who assigned the value “1” to answers with a clitic and “0” when the clitic was absent. For a small proportion of trials participants gave an answer that did not provide a context where a clitic could have been used, such as the word of affirmation *da* or use of verbs where the intended object functioned as an experiencer subject, such as with the verb *place* ‘like’ or verbs in the passive voice. For this reason, 17 items (4.7%) for the Italian

TABLE 5 Written Elicitation task (WET) results from Romanian monolinguals, Romanians in English speaking countries and Romanians in Italy.

	Effects of specificity and discourse on use of clitic			
	Italian monolings	Romanian monolings	L2 English	L2 Italian
Predictor	Estimate(SE), <i>p</i> -value	Estimate(SE), <i>p</i> -value	Estimate(SE), <i>p</i> -value	Estimate(SE), <i>p</i> -value
Intercept	0.65 (0.66)	−0.27 (3.43)	0.63 (0.98)	0.60 (0.93)
Anaphoricity Top-Foc	9.08 (1.95), <i>p</i> < 0.001	9.11 (19.07)	−0.53 (6.82)	0.66 (1.42)
Specificity Spec-NonSpec	2.25 (1.05), <i>p</i> = 0.033	48.35 (12.03), <i>p</i> < 0.001	7.4 (1.94), <i>p</i> < 0.001	9.39 (1.95), <i>p</i> < 0.001
Anaphoricity Top-Foc: Specificity Spec-NonSpec	−0.18 (1.91)	−2.88 (18.75)	−5.53 (3.55)	−12.65 (4.03), <i>p</i> < 0.01

Showing effects of specificity and discourse on use of clitic: estimate for each predictor with SE in parentheses. Significant effects are shown in bold, and *p*-values are provided.

monolinguals, 14 items (4.1%) for the Romanian monolinguals and 46 items (4.9%) for the attriter groups were excluded from the analysis.

Figure 4 shows the proportion of clitics used by Italian monolinguals, Romanian monolinguals, Romanian immigrants to anglophone countries and Romanian immigrants to Italy, and Table 5 shows the results of binary logistic regression models. Random effects for Participant and Item were included where possible, aiming for a maximal model that allowed convergence (Barr et al., 2013).

As can be observed in Figure 4, Italian monolinguals used clitics in both topic conditions, regardless of specificity. Speakers in all three native Romanian groups rarely used a clitic when the fronted object was non-specific. Results from logistic regression models confirm that the odds of using a clitic are significantly greater in the specific conditions than in the no-clitic conditions. Additionally, the L2 Italian group shows a significant interaction between specificity and discourse. In addition to clitics being used more when the fronted object is specific, clitics were also used more in the specific topic condition than the specific focus condition.

6.2.3 Interim conclusion

The results from both the Acceptability Judgment and the Written Elicitation task show no significant effect of attrition for the Romanian native speakers who immigrated to an anglophone country. These participants fully retained the specificity feature associated with CLLD in Romanian. Romanians residing in Italy, however, also accepted and used clitics when the fronted object was [+specific], but discourse anaphoricity, the property that is associated with CLLD in Italian, seems to interfere: in the Acceptability Judgment task the group result showed no difference between clitic and no-clitic sentences with fronted non-specific topics, and in the Written Elicitation task, clitics were not used consistently with specific foci. The differences in performance patterns between the L2 English and the L2 Italian group are in line with the prediction that attrition occurs when individuals are exposed to a syntactic structure that exists in both the L1 and L2 but are nonetheless distinct, as these L2 features are susceptible to competition. Since attrition is typically characterized by individual variation and to further examine whether the group results discussed in Section 6 are driven by specific individuals, we examine effects of language dominance and age of onset of L2 acquisition (using the criteria from Table 3) in Section 7. This section focuses only on the data from the L2 Italian speakers, as L1-L2 differences were only found in this group.

7 Effects of age of arrival and language dominance

Attrition is typically characterized by individual variation and the extent at which attrition occurs has been argued to be modulated by (a combination of) various factors, including age of immigration and relative language use.

The most prominent non-linguistic factor for attrition is age of onset of bilingualism: attrition has been shown to be quite common for speakers who immigrate before the onset of puberty (roughly between the ages of 8 and 13 years; Köpke and Schmid, 2004; Pallier, 2007; Bylund, 2009), in particular for morpho-syntactic phenomena (e.g., Montrul, 2008; Montrul et al., 2014). The grammars of post-puberty bilinguals are much more stable and restructuring is argued to be fairly rare (Köpke and Schmid, 2004). Crucially, however, very few studies have looked at the grammars of adolescents, as most research has focused on the grammars of speakers who became bilingual in childhood (heritage language acquisition) or in post-puberty bilinguals (attrition studies; Schmid and Köpke, 2017a). Since the age of arrival (AoA) of the L2 Italian group ranged from age 11 to 40, we have the data to examine potential effects of AoA and to examine whether attrition is more likely in those who immigrated during adolescence.

The results of the Acceptability Judgment task showed that Romanians in Italy differed from the other native Romanian groups in how they distinguished between clitic and no-clitic sentences in the non-specific topic condition. This is also the discourse context for which the AvA predicts seeing effects of attrition, as the use of clitics in Italian in this condition could be transferred to Romanian. Figure 3 showed that Romanians in Italy rated clitic and no-clitic sentences as equally acceptable with non-specific topics, while Italian monolinguals preferred clitic sentences and the other Romanian native groups preferred no-clitic sentences in this condition. It is, however, possible that this group effect is driven by some speakers rating no-clitic sentences and others rating clitic sentences as more acceptable. To visualize individual differences and to examine trends in the data on how age of arrival may have affected the ratings of clitic versus no-clitic sentences in the non-specific topic condition, we calculated a new composite dependent variable showing how each individual distinguished between clitic and no-clitic sentences in this condition.⁴

⁴ I would like to thank Bradley Hoot for his suggestion to visualize potential effects of AoA this way.

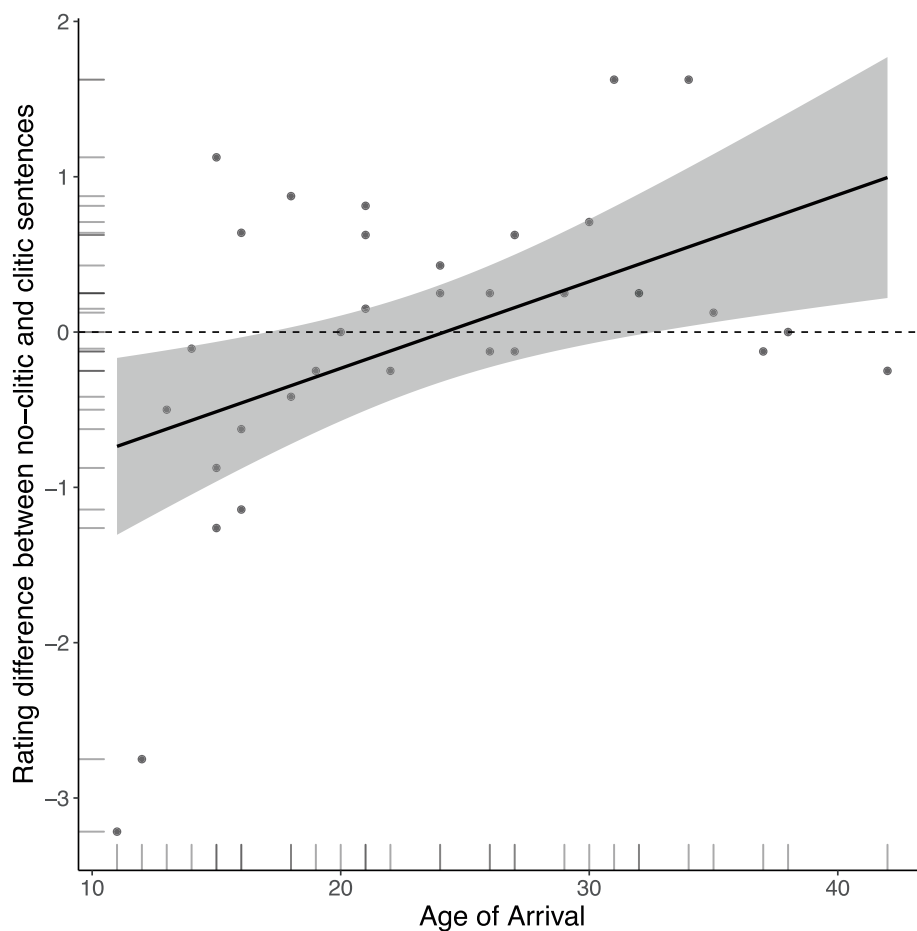


FIGURE 5
Differences in judgments between no-clitic and clitic sentences in the non-specific topic condition, showing results from Romanian native speakers residing in Italy.

We subtracted each participant's average rating for sentences without clitics from those with clitics. A negative value therefore indicates a higher rating for clitic sentences and a positive value a higher rating for no-clitic sentences. [Figure 5](#) shows how the difference scores change as a function of age of arrival. The increasing line in the scatterplot, whereby earlier arrivals tend to rate clitic sentences as more acceptable and later arrivals tend to rate no-clitic as more acceptable, suggests that AoA affects acceptability judgments of clitic and no-clitic sentences with fronted non-specific topics. We furthermore observe that more speakers with an earlier AoA, especially those who migrated to Italy before their mid-twenties, had a difference score below 0 (below the dashed line), meaning that clitic sentences were preferred for these speakers.

[Figure 6](#) shows individual differences in the use of clitics in the Written Elicitation task, plotted against age of arrival. We are showing the results of two conditions where Romanian and Italian differ and where more variability is expected, namely the non-specific topic and specific focus conditions. For sentences with non-specific topics, a high use of clitics suggests attrition, while for sentences with specific foci, a low use of clitics suggests attrition. The results from the non-specific topic condition show that most participants never used a clitic in this condition. However, note that there are two early arrivals (their AoA = 11 and 12) who consistently used clitics in this condition.

Their results are likely driving the AoA effect suggested by the trend line in the scatterplot. For the specific focus condition, the relatively flat line suggests that there is no AoA effect for the specific focus condition. In fact, there is much more variability across individuals in their use of clitics in this condition.

It is important to point out that the AoA effects we observed for the non-specific topic condition can also be due to a difference in language dominance between earlier and later migrants. In fact, the literature on attrition reports a relationship between attrition and language dominance and "... even if a reversal in language dominance is not necessarily followed by attrition, it is most likely that attrition is preceded by such a reversal..." ([Köpke and Schmid, 2004](#), p. 12). Although language dominance is often defined as the relative proficiency in each of a bilingual's languages, it has been proposed for adult bilinguals that language dominance may be independent from language proficiency ([Gertken et al., 2014](#)) and also that dominance is a complex interaction between proficiency and input components (e.g., [Montrul, 2015](#)). Because of this and due to the lack of a standardized measure to determine language dominance using the questionnaire we adopted, we define language dominance as a mix between proficiency and exposure/language use criteria. As discussed in Section 6.1, our background questionnaire contained questions on language proficiency, external language use and internal language use. [Figure 7](#)

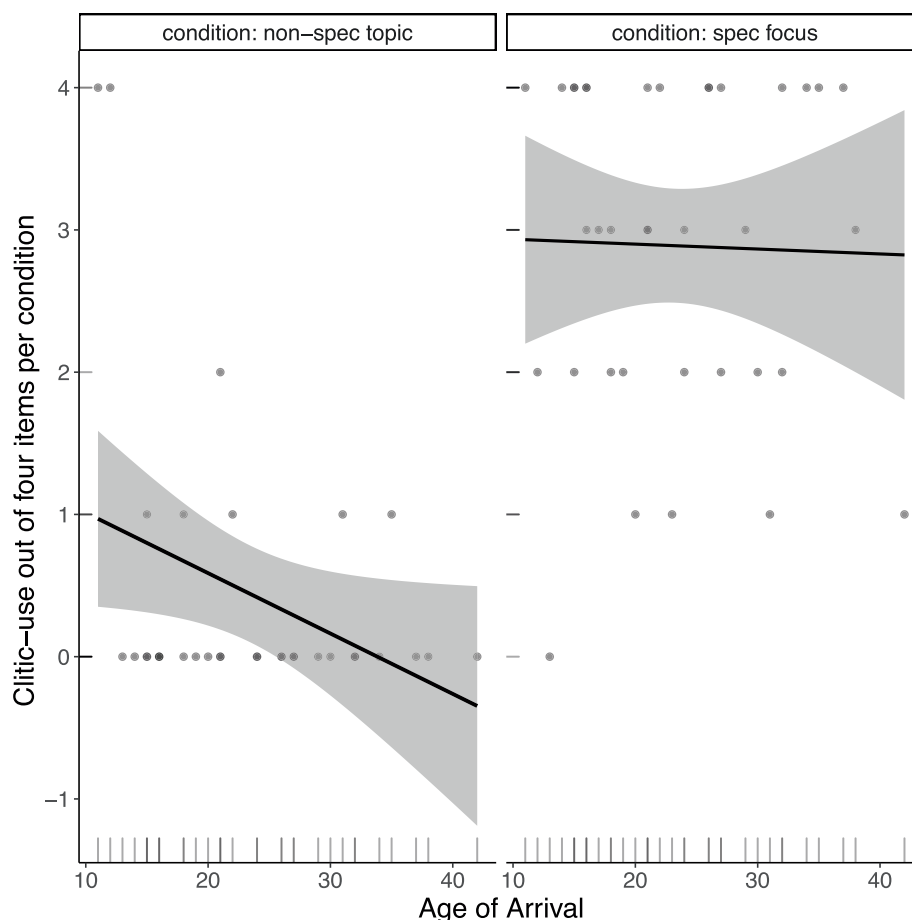


FIGURE 6

Clitics used by individual participants in the non-specific topic and specific focus conditions, showing results from Romanian native speakers residing in Italy.

shows the individual average scores for questions on each of these three factors and its correlation with AoA. Based on the five-point scale used in our background questionnaire, an average score lower than 3 indicates a higher proficiency and use of Italian, while a score higher than 3 indicates a higher proficiency and use of Romanian. The plot shows that those individuals who arrived in Italy before their mid-twenties tended to use Italian more than Romanian (“External language use”), while the opposite holds for those who immigrated post-adolescence. Similar AoA effects are observed for “Internal language use” and “Relative language proficiency,” where those who immigrated at a later age use Romanian more than Italian and judge themselves as more proficient in their L1. Note furthermore that for “External language use” about half of our participants use Italian more while the other half uses Romanian more. However, for “Internal language use” and “Relative language proficiency,” more participants showed an average rating above 3 (above the dashed line), suggesting that their linguistic and cultural affiliation is more strongly connected with Romanian. In Section 8.1 we will elaborate on why many of our participants may be more connected to their Romanian identity and claim to be more proficient in Romanian than in Italian despite having lived in Italy for 10+ years.

To summarize, a closer look into the performance of the L2 Italian group revealed that the group results in our experimental tasks were mainly driven by speakers who immigrated to Italy during adolescence

and in their early twenties and who were more dominant speakers of Italian. Recall that the AvA predicts grammatical attrition to occur in the form of L2 options being added to the existing L1 options, predicting an increased use of clitics with non-specific topics, as this is where clitics are used in Italian but not in Romanian. For the AJ task we found that earlier arrivals tend to rate clitic sentences as more acceptable than no-clitic sentences (like Italian monolinguals), while the opposite tends to hold for later arrivals. In the WET, however, use of clitics with non-specific topics was only found for two participants. We will return to possible task effects to explain this difference in Section 8.2.

8 Discussion

This paper examined the use of CLLD among native Romanian speakers in an L2 Italian or L2 English immersion context to test the role of L2 acquisition on native language attrition. The main findings from the Acceptability Judgment and the Written Elicitation task can be summarized as follows: the acquisition of an L2 property with similar behavior in the L1 and L2 (that are nevertheless different), but not a reduction of L1 input alone, causes attrition, as only Romanians in Italy behaved significantly differently from the Romanian monolingual control group. A further examination of language-external factors revealed that Romanians who immigrated during adolescence or in their early twenties

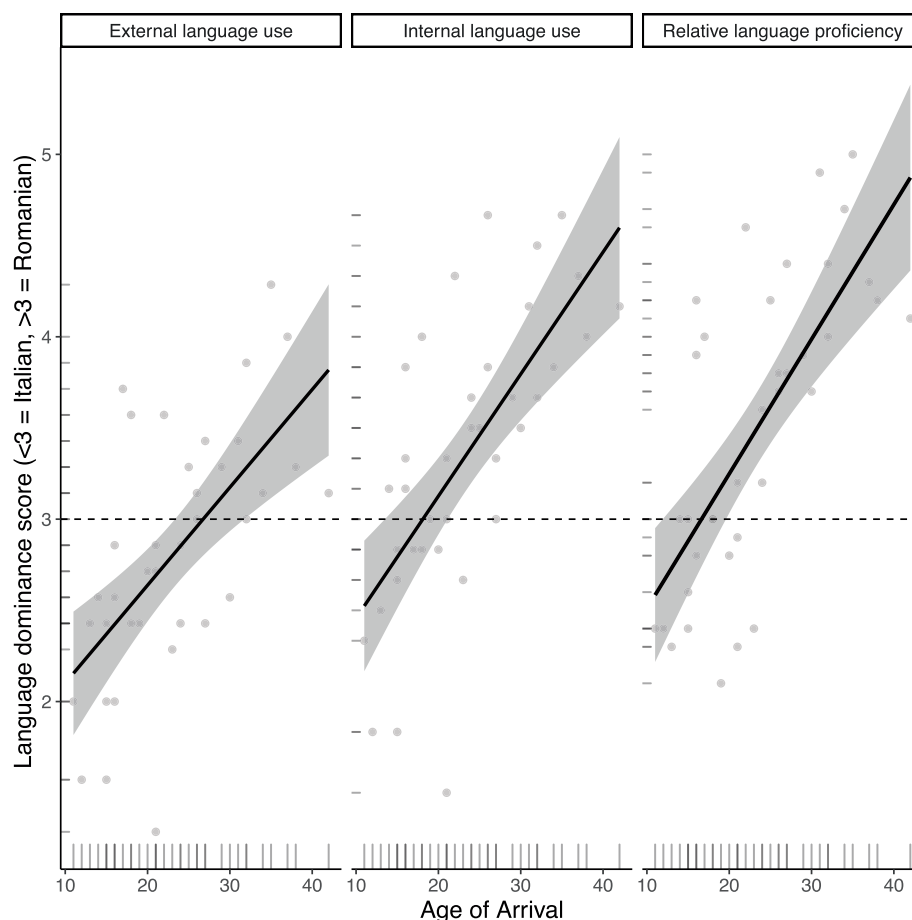


FIGURE 7

Correlation between age of arrival and language dominance scores for Romanians residing in Italy.

were more susceptible to L2 transfer. In the Acceptability Judgment task, they accepted clitics in all discourse contexts where clitics are used in either their L2 Italian or their L1 Romanian and in the Written Elicitation task they applied clitics less deterministically. Importantly, however, participants with an earlier age of immigration typically also reported higher levels of L2 proficiency and L2 language use, which may have been the causing factor of attrition.

8.1 Effects of age of onset and quality and quantity of exposure

In this section, I briefly elaborate on whether someone's language is maintained or attrited depends on a number of variables that are independent from maturational effects, such as the quantity and quality of input and the level of engagement with both the L1 and the L2. I furthermore discuss specific characteristics of the Romanian community in Italy and the group of subjects who participated in this study to discuss why language maintenance was fairly high in this study.

One question is why attrition effects tend to be higher in speakers who arrive in the L2 country during adolescence. It has been argued that teens are typically at an age where many start post-secondary education in the L2 and form social relationships outside of their heritage community. They have broader friend circles and use social media and

social networking more, increasing the quality and quantity of L2 input, which in turn increases L2 proficiency and language use [Anderson and Jiang, 2018; Roehl and Stewart, 2018, as cited in Miller and Rothman (2020), who also found a difference between participants who migrated before and after the age of 22]. In consequence, they may be using the L2 to a greater extent than the L1 and to a greater extent than speakers who immigrated later in life. Examining potential effects of schooling in the L2 in our participant pool, it is noteworthy that eight out of 15 participants who arrived before the age of 20 reported high school as their highest level of education. Four of them arrived between the ages of 18 and 20 and therefore never received schooling in Italian while the other four did receive some schooling in Italian (one participant moved at age 14, two at age 16 and one at age 17). In other words, not all participants with an AoA during adolescence have necessarily received schooling in their L2. Additionally, the number of participants who reported high school as their highest level of education is quite high (16 out of the total 37 Romanians in Italy) and therefore many of our participants did not receive any post-secondary training in Italian either. Furthermore, most of these 16 participants are currently working as housewives, drivers or seasonal workers, jobs that likely do not require a near-native level of proficiency in Italian. When the need to speak the L2 is low, L2 interference is naturally low as well.

Additionally, rates of attrition are likely low for our participants as there are many opportunities to speak Romanian, both inside and outside

their homes, and relatively few of our participants became dominant speakers of Italian. Maintaining close contacts with the L1 speaking community reinforces the L1 language system. The Romanian diaspora is the fifth largest in the world and one third of all Romanian emigrants (over 1 million) are living in Italy (OECD, 2019), a country with 55 million citizens. Most Romanians in Italy are residing in metropolitan cities such as Rome, Turin and Milan and the industrial areas of northern Italy, of which the province of Veneto is one of the largest (Stocchiero, 2002). The majority of participants in our study are family friends or members of the church community of one of the Research Assistants, who all reside in Veneto. This province has around 126,344 Romanian-born citizens (ISTAT, 2023). Our participants are part of a community of Romanian speakers with relatively high enclosure and likely have ample opportunity to maintain their native language. Many Romanians are able to successfully emigrate because of family members already residing in Italy, helping newcomers with accommodation and employment (Stocchiero, 2002). In fact, the 60 Romanians tested in our study almost all report that they speak Romanian daily (with four participants stating they speak Romanian only weekly). Additionally, given the geographical closeness between Romania and Italy, Romanians in Italy typically have ample opportunities to visit their home country. In fact, 25 out of 37 participants in Italy report to visit Romania at least once or twice per year. Furthermore, all but two participants in this group report to frequently be in touch with Romanian relatives and friends in Romania, with whom they exclusively speak in Romanian. Remarkably, the left plot in Figure 7 also showed that about half of our participants use Romanian on a daily basis at least as much as they use Italian, despite living in Italy.

Even though language maintenance levels are high for the speakers of the Romanian community in Italy with whom we conducted the experiments, we do find effects of L2 induced attrition and clear cases of individuals in our study for whom a shift in language dominance to the L2 caused restructuring of the L1 grammar. For example, two of our participants who immigrated at the ages of 11 and 12 showed complete L2 convergence in the non-specific topic condition. These individuals immigrated during high school age and obtained a university degree in Italy. They are now working in an Italian-speaking environment and use Romanian mainly with their families. These findings are in line with previous studies that showed that speakers who use their L1 mainly in informal contexts experienced higher attrition levels (Schmid, 2007; De Leeuw et al., 2010; Schmid and Dusseldorp, 2010; de Leeuw et al., 2012; Yilmaz and Schmid, 2012). Our findings also show that individuals who arrived before their early twenties and who have become dominant speakers of the L2 show the greatest extent of non-monolingual-like use of CLLD. Experimental research on attrition naturally tends to focus on diaspora communities for which a sufficient number of participants can be found. However, these speakers naturally may have more opportunity for language maintenance, leading researchers to conclude that attrition levels are low. This observation encourages future attrition research to keep the L2 community status into consideration as a factor contributing to L1 attrition.

8.2 Grammatical attrition and task effects

In Section 2 we discussed the question whether performance differences between attriters and non-attriters are the result of a conflict between the two grammars at the level of sentence processing alone or

whether they constitute a difference in knowledge representation. It is relevant to observe that both approaches discuss the notion of “optionality,” where attriters accept both an interpretation/structure allowed in the L1 as well as the equivalent from the L2. This is in fact the most attested form of attrition for morpho-syntactic phenomena (as opposed to only using/accepting the L2 option). For processing theories like the Interface Hypothesis (Sorace, 2011), optionality means that L1 attriters are less deterministic in their choices. For the Attrition via Acquisition model (Hicks and Domínguez, 2020), a theory of linguistic representation, optionality between L1-L2 structures suggests that both options are part of the grammar and can therefore be grammatical for an attrited speaker. Since all experiments are measures of performance, it is often hard to disentangle which differences between attriters and control groups represent a qualitatively different grammar and which are temporary effects of crosslinguistic influence from the L2 grammar. For the results of the current study, however, I argue that the reported findings indicate a clear modification of the L1 grammar. Namely, for the conditions where the L1 and the L2 differ, which is with non-specific topics and with specific foci, one may expect attriters to accept both clitic and no-clitic sentences, as each can be parsed by either one of the options available in the L1 or the L2. For left dislocation with non-specific topics, for example, Italian does use a clitic but Romanian does not, and an attriter who has both options available is then expected to accept both the clitic and the no-clitic sentences. However, speakers in the attrited group (the L2 Italian speakers in Group 1) rate clitic sentences as more acceptable than no-clitic sentences, suggesting that the L2 grammar replaced the L1 options for that particular condition. The additional option of CLLD in Italian was added to the L1 Romanian grammar of attrited speakers and within that specific context the option of adding a clitic replaced the availability of no-clitic sentences.

One surprising finding is the fact that we did not find an increased use of clitics with non-specific topics in the Written Elicitation task for more speakers who rated clitics as acceptable in this condition in the Acceptability Judgment task. This difference in outcome may be due to the difference in task demand. While no time constraint was applied to either of the offline tasks, the Acceptability Judgment task naturally requires participants to respond faster than the Written Elicitation task. Since participants cannot hear the question-and-answer pair again, this task more naturally taps into the speaker's intuition. The Written Elicitation task is more meta-linguistic in nature and participants may have applied formal reasoning strategies when giving their responses, if they become aware that clitics are used only with specific objects (shown in the experiment with definite articles). As linguists we use the intuitive judgments of speakers to describe their mental grammars. Since the Acceptability Judgment task more clearly taps into intuitive knowledge, the results from this task may be a better representation of the speakers' grammars. In future studies, it is important to include other tasks, such as a spoken elicitation task, where participants cannot apply formal reasoning strategies due to communication pressure, or a self-paced reading experiment to measure in which discourse contexts participants expect a clitic.

8.3 Connecting L1 attrition to L2 acquisition

Research on L1 attrition has started to become more systematically connected to research on L2 acquisition and resemblances have been

observed between the performances of near-native speakers and attriters (see for example Sorace, 1999, 2000a, b, 2003 on the use of overt pronouns in pro-drop languages). Montrul (2020, p. 214) furthermore states that “Fossilization ... could be seen as the opposite of attrition because despite optimal input, the inference module seems not to be engaged or fails to become engaged to change grammatical representations.” More clearly, what fossilization and attrition have in common is that the perceptual representations of the grammars in the minds of speakers are not compatible with the input these learners are exposed to. A crucial difference between L2 end-state grammars and L1 attrited grammars is that while few L2ers reach full target-like competence in the L2, L1 grammars rarely change. Hicks and Domínguez (2020) explain this so-called paradox by suggesting that speakers continue to process input for the purpose of acquisition, as long as there is some new form of input that is structurally different from the existing mental L1 grammar. In the case of CLLD, there is coactivation with a competing L2 system due to the structural overlap between Romanian and Italian.

Since the Attrition via Acquisition model characterizes attrition as a potential outcome of acquiring another language, attrition is predicted to be possible only when L2 acquisition has occurred. The current study did not examine the L2 grammars of Romanians in Italy. This is, however, the exact population tested in Smeets (2023), who reports on the results of English and Romanian L2 speakers of Italian and English and Italian L2 speakers of Romanian at two levels of proficiency: high intermediate/advanced and near-native. The study used the same Acceptability Judgment and Written Elicitation task and only the results from the Italian-Romanian speakers are relevant here, as no transfer is possible for L1 English groups. The acquisition task involved the reorganization of grammatical features from the transferred L1 grammar to match those of the L2 input. Specifically, for complete acquisition, Romanian L2 learners of Italian have to remove the [+specific] feature and acquire the [+anaphor] feature, which proves to be successful if they start using clitics with fronted non-specific topics and stop using clitics with specific foci. Italian L2 learners of Romanian have to do the opposite by removing the [+anaphor] feature and replacing it with a [+specific] feature. The results from both tasks showed that at the near-native stages of L2 proficiency, but not earlier, speakers in both learning directions were able to broaden the contexts that use a clitic in the L2 (grammatical expansion), but L1 pre-emption difficulties were attested as well. That is, Italian L2 learners of Romanian correctly acquired the use of clitics with fronted specific foci and Romanian L2 learners of Italian did so with fronted non-specific topics. However, neither group rejected or stopped using clitics in discourses where they are allowed in the L1 but not in the L2: Romanian L2 learners of Italian continued to use clitics with specific foci and Italian L2 learners of Romanian used clitics with non-specific topics. Thus, the study found that the L2 options were added to the options transferred from the L1. Like the first-generation immigrants in the current study, most near-native L2ers in Smeets (2023) had been acquiring Italian (or Romanian) for 10+ years and had been living in Italy (or Romania). Their L2 grammars had likely fossilized.

Similar findings are reported in Gürel (2002, 2007), the first study to directly compare the knowledge representation of L1 attriters to L2 near-native speakers, which also found that in both end-state and attrited grammars, the options from the L1 and the L2 are merged. She compared the Turkish of Turkish speakers living in North America (end-state L2

English speakers) discussed in Section 2 to the Turkish of English speakers living in Turkey (end-state L2 Turkish speakers) on their interpretation of the binding properties of the Turkish pronoun *o*. Recall that English pronouns *him/her* can be bound within a larger domain in English than Turkish *o*. Gürel (2002, 2007) ‘Set-Theoretic Language Attrition Model’ suggests that attrition is most favorable when the L2 allows a superset of the interpretations available in the L1 and therefore when the L1 is more restrictive. The results indeed showed that both the attriters as well as the near-native speakers of Turkish added the interpretation from English pronouns *him/her* to that of their Turkish grammars. In other words, the L1 affected near-native L2 grammars similarly to how L2 grammar affects L1 attrited grammars. In both L2 acquisition and L1 attrition, the input data from one of the grammars provides evidence for this additional option. Note furthermore that the premise of the AvA is similar to that of the ‘Set-Theoretic Language Attrition Model’, as it suggests that grammatical attrition is disfavoured if it involves losing an option from the L1. In fact, Hicks et al. (2024) recently found no attrition in Spanish immigrants to the United Kingdom on their aspectual interpretations of the Spanish present tense and attribute this to the fact that while Spanish allows both an ongoing and habitual interpretation, English only allows the habitual interpretation. Exposure to L2 English can therefore not add options to the L1 grammar that do not already exist in the L1. In the current study we also found that Romanians in Italy can add the use of clitics with non-specific topics, as they are also used in this discourse context in Italian, but attrition does not involve the loss of clitics with specific foci.

9 Conclusion

This study examined whether the interpretational properties of Clitic Left Dislocation in Romanian first generation immigrants are subject to grammatical attrition. We discussed that continued activation of a similar property in the L2 grammar can impact parsing strategies of the L1 to an extent that the L2 parse can (permanently) change the grammatical representation of that syntactic construction. We focused on a discourse-syntactic phenomenon called Clitic Left Dislocation and examined whether grammatical attrition, in the form of L2 options or morpho-syntactic properties being added to the L1 grammar, can occur in the L1 linguistic competence of native Romanian speakers who are late L2 learners of either English or Italian. Results from the Acceptability Judgment task most specifically showed attrition in Romanian native speakers who moved to Italy during adolescence and who were likely most integrated in the L2 Italian community. While Romanian L2 speakers of English and Romanian immigrants to Italy who moved after adolescence did not differ from Romanians in Romania, earlier immigrants to Italy allowed clitics in Romanian also in the discourses where they are allowed in Italian but not in Romanian. Our findings contribute to an increasing body of literature showing that L1 attriters and L2 learners can end up with very similar grammars and confirm the importance of studying second language acquisition and L1 loss within a broader picture of bilingual development.

In keeping with the research topics “Experimental Approaches to the Acquisition of Information Structure” the current study drew on linguistic data from underrepresented populations, as CLLD has not previously been studied in attrition research, as well as underrepresented languages, as the L2 acquisition of discourse constraints on CLLD has previously only been studied using English native speakers acquiring an L2 with

CLLD, mainly Spanish. We furthermore examined the research subjects from different methodological perspectives.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://osf.io/5xmfw/?view_only=3c03f17e284b409fbc328316d3c150c0.

Ethics statement

The studies involving humans were approved by Office of Research Ethics, York University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

LS: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

References

- Altenberg, E. P. (1991). "Assessing first language vulnerability to attrition" in *First language attrition*. eds. H. W. Seliger and R. M. Vago (New York, NY: Cambridge University Press), 189–206.
- Anderson, M., and Jiang, J. (2018). Teens, social media & technology. *Pew Res. Center* 31:2018.
- Avram, L., and Coene, M. (1999). Null objects and accusative clitics in Romanian. *Bucharest Working Papers in Linguistics* IV, 88–106.
- Baker, L. (2024). The role of linguistic input in adult grammars: Modelling L1 morphosyntactic attrition. Ph.D. Thesis. University of Southampton.
- Barr, D., Levy, R., Scheepers, C., and Tilly, H. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *J. Mem. Lang.* 68, 255–278. doi: 10.1016/j.jml.2012.11.001
- 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
- Bley-Vroman, R. (1983). The comparative fallacy in interlanguage studies: the case of systematicity. *Lang. Learn.* 33, 1–17. doi: 10.1111/j.1467-1770.1983.tb00983.x
- Blumenfeld, H. K., and Marian, V. (2007). Constraints on parallel activation in bilingual spoken language processing: examining proficiency and lexical status using eye-tracking. *Lang. Cogn. Process.* 22, 633–660. doi: 10.1080/01690960601000746
- Büring, D. (2002). On D-trees, beans, and B-accent. *Unpublished manuscript*, Los Angeles: University of California.
- Bylund, E. (2009). Maturational constraints and first language attrition. *Language Learning*. 59, 687–715. doi: 10.1111/j.1467-9922.2009.00521.x
- Cardinaletti, A., and Starke, M. (1994). The typology of structural deficiency. A case study of the three classes of pronouns. Cambridge (MA): Massachusetts Institute of Technology.
- Chamorro, G., and Sorace, A. (2019). "The Interface hypothesis as a framework for studying L1 attrition" in *The Oxford handbook of language attrition*. eds. M. S. Schmid and B. Köpcke (Oxford: Oxford University Press), 25–35.
- Chamorro, G., Sorace, A., and Sturt, P. (2016). What is the source of L1 attrition? The effect of recent L1 re-exposure on Spanish speakers under L1 attrition. *Biling. Lang. Cogn.* 19, 520–532. doi: 10.1017/S1366728915000152
- Chamorro, G., Sturt, P., and Sorace, A. (2016). Selectivity in L1 attrition: differential object marking in Spanish near-native speakers of English. *J. Psycholinguist. Res.* 45, 697–715. doi: 10.1007/s10936-015-9372-4
- Chomsky, N. (2000). "Minimalist inquiries: the framework (MITOPL 15)" in *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*. eds. M. R. Michaels, D. Uriagereka and J. Martin (Cambridge, MA: MIT Press), 89–155.
- Chomsky, N. (2001). "Derivation by phase" in *Ken Hale: A life in linguistics*. ed. M. Kenstowicz (Cambridge: MIT Press), 1–52.
- Christensen, R. H. B. (2014). Analysis of ordinal data with cumulative link models: Estimation with the r-package ordinal. Available at: https://www.people.vcu.edu/~dbandyop/BIOS625/CLM_R.pdf [Accessed February 2023].
- Cornilescu, A., and Dobrovie-Sorin, C. (2008). "Clitic doubling, complex heads, and interarboreal operations" in *Clitic doubling in the Balkan languages*. eds. L. Tasmowski and D. Kallulli (Amsterdam: John Benjamins Publishing Company), 289–219.
- De Leeuw, E., Schmid, M. S., and Mennen, I. (2010). The effects of contact on native language pronunciation in an L2 migrant setting. *Biling. Lang. Cognit.* 13, 33–40.
- De Leeuw, E., Mennen, I., and Scobbie, J. M. (2012). Singing a different tune in your native language: first language attrition of prosody. *Int. J. Biling.* 16, 101–116. doi: 10.1177/1367006911405576
- Dominguez, L. (2013). Understanding interfaces: Second language acquisition and native language attrition of Spanish subject realization and word order variation (language acquisition and language disorders, 55). Amsterdam, NL: John Benjamins Publishing Company. Amsterdam, NL.
- Dominguez, L., and Hicks, G. (2016). "Synchronic change in a multidialectal community: evidence from Spanish null and postverbal subjects" in *Inquiries in Hispanic linguistics: From theory to empirical evidence*. eds. A. Cuza, L. Czerwionka and D. Olson (Amsterdam: John Benjamins Publishing Company), 53–72.
- Gertken, L. M., Amengual, M., and Birdsong, D. (2014). "Assessing language dominance with the bilingual language profile" in *Measuring L2 proficiency: Perspectives from SLA*. eds. P. Leclercq, A. Edmonds and H. Hilton (Bristol: Multilingual Matters), 208–225.
- Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., and Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Mem. Cogn.* 33, 1220–1234. doi: 10.3758/BF03193224

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. The Social Sciences and Humanities Research Council (P2021-3842).

Acknowledgments

I would like to thank Ileana Grama for her help with stimuli design, Anda Neagu for her help with implementing the experiments and coding of responses, and Alexandra Gabriela Cretan and Alexandra Stoica for participant recruitment and testing.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Grimshaw, J., and Samek-Lodovici, V. (1998). "Optimal subjects and subject universals" in *Is the best good enough*. eds. P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis and D. Pesetsky (Cambridge, MA: MIT Press), 193–220.
- Gürel, A. (2002). Linguistic characteristics of second language acquisition and first language attrition: Overt versus null pronouns [doctoral dissertation]. [Montreal, (CA)], McGill University.
- Gürel, A. (2004). "Attrition in L1 competence: the case of Turkish" in *First language attrition: Interdisciplinary perspectives on methodological issues*. eds. M. S. Schmid, M. Keijzer and L. Weilemar (Amsterdam: John Benjamins Publishing Company), 225–242.
- Gürel, A. (2007). "(psycho)linguistic determinants of L1 attrition" in *Language attrition: A theoretical perspective*. eds. B. Köpke, M. S. Schmid, M. Keijzer and S. Dostert (Amsterdam: John Benjamins Publishing Company), 99–120.
- Gürel, A., and Yılmaz, G. (2011). Restructuring in the L1 Turkish grammar: effects of L2 English and L2 Dutch. *Language. Interaction Acquisition* 2, 221–250. doi: 10.1075/lia.2.2.03gur
- Hicks, G., and Domínguez, L. (2020). A model for L1 grammatical attrition. *Second. Lang. Res.* 36, 143–165. doi: 10.1177/0267658319862011
- Hicks, G., Laura Domínguez, L., Jamieson, E., and Schmid, M. S. (2024). L1 grammatical attrition in late Spanish-English bilinguals in the UK: aspectual interpretations of present tense in Spanish. *Lang. Learn. J.* 52, 145–158. doi: 10.1080/09571736.2023.2293024
- ISTAT, (2023). Available at: <https://www.istat.it/it/> (Accessed: 15 June 2024).
- Iverson, M. (2012). Advanced language attrition of Spanish in contact with Brazilian Portuguese [Doctoral dissertation]. [Iowa City (IA)]: University of Iowa.
- Jitcă, D., Apopei, V., Păduraru, O., and Marușca, S. (2015). "Transcription of Romanian intonation" in *Intonation in romance*. eds. S. Frota and P. Prieto (Oxford: Oxford University Press), 285–316.
- Kaltsa, M., Tsimpli, I. M., and Rothman, J. (2015). Exploring the source of differences and similarities in L1 attrition and heritage speaker competence: evidence from pronominal resolution. *Lingua* 164, 266–288. doi: 10.1016/j.lingua.2015.06.002
- Kasparian, K. (2015). The case of the non-native-like first-language: Neurophysiological investigations of first-language attrition and second-language processing [Doctoral dissertation]. [Montreal (CA)]: McGill University.
- Köpke, B. (1999). L'attrition de la première langue chez le bilingue tardif: Implications pour l'étude psycholinguistique du bilinguisme [first language attrition in late bilinguals: Implications for the psycholinguistic research of bilingualism.]. [Doctoral dissertation]. [Toulouse (FR)]: Université de Toulouse-Le Mirail.
- Köpke, B. (2002). "Activation thresholds and non-pathological first language attrition" in *Advances in the neurolinguistics of bilingualism*. ed. F. Fabbro (New York, NY: Routledge), 119–142.
- Köpke, B., and Schmid, M. S. (2004). "Language attrition: the next phase" in *First language attrition: Interdisciplinary perspectives on methodological issues*. eds. M. Schmid, B. Köpke, M. Keijzer and L. Weilemar (Amsterdam: John Benjamins Publishing Company), 1–43.
- Lardiere, D. (2008). "Feature assembly in second language acquisition" in *The role of formal features in second language acquisition*. eds. J. Liceras, H. Zobl and H. Goodluck (New York, NY: Lawrence Erlbaum), 106–140.
- Lidz, J., and Gagliardi, A. (2015). How nature meets nurture: universal grammar and statistical learning. *Annual Rev. Linguistics*. 1, 333–353. doi: 10.1146/annurev-linguist-030514-125236
- López, L. (2009). A derivational syntax for information structure. Oxford: Oxford University Press.
- Marian, V., and Spivey, M. (2003). Competing activation in bilingual language processing: within- and between-language competition. *Biling. Lang. Cogn.* 6, 97–115. doi: 10.1017/S1366728903001068
- Miličević, M., and Kraš, T. (2017). Translation between L2 acquisition and L1 attrition: anaphora resolution in Italian by English-Italian trainee translators. *Appl. Linguist.* 38, 21–42. doi: 10.1093/applin/amu070
- Miller, D., and Rothman, J. (2020). You win some, you lose some: comprehension and event-related potential evidence for L1 attrition. *Biling. Lang. Cogn.* 23, 869–883. doi: 10.1017/S1366728919000737
- Montrul, S. (2008). *Incomplete Acquisition in Bilingualism: Re-examining the age factor*. Amsterdam: John Benjamins Publishing Company.
- Montrul, S. (2015). *The acquisition of heritage languages*. Cambridge: Cambridge University Press.
- Montrul, S. (2020). Attrition, addition, and age. *Second. Lang. Res.* 36, 213–217. doi: 10.1177/0267658319889675
- Montrul, S., Davidson, J., De la Fuente, I., and Foote, R. (2014). Early language experience facilitates the processing of gender agreement in Spanish heritage speakers. *Biling. Lang. Cogn.* 17, 118–138. doi: 10.1017/S1366728913000114
- OECD (2019). *Talent abroad: A review of Romanian Emigrants, Talent Abroad*. Paris: OECD Publishing.
- 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 Publishing Company), 155–168.
- Paradis, J. (2007). Bilingual children with specific language impairment: theoretical and applied issues. *Appl. Psycholinguist.* 28, 551–564. doi: 10.1017/S0142716407070300
- Pelc, L. (2001). L1 lexical, morphological and morphosyntactic attrition in Greek-English bilinguals [Doctoral dissertation]. [New York (NY)]: City University of New York.
- Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation* [Doctoral dissertation]. [Cambridge (MA)]: Massachusetts Institute of Technology.
- Putnam, M. T., and Sánchez, L. (2013). What's so incomplete about incomplete acquisition? A prolegomenon to modeling heritage language grammars. *Linguistic Approaches to Bilingualism*. 3, 478–508. doi: 10.1075/lab.3.4.04put
- Roehl, A., and Stewart, A. H. (2018). Impact of social networking sites and digital applications upon teens. *J. Fam. Consum. Sci.* 110, 37–42. doi: 10.14307/JFCS110.2.37
- Rothman, J., and Slabakova, R. (2011). The mind-context divide: on acquisition at the linguistic interfaces. *Lingua* 121, 568–576. doi: 10.1016/j.lingua.2011.01.003
- Schmid, M. S. (2007). "The role of L1 use for L1 attrition" in *Language attrition. Theoretical perspectives*. eds. B. Köpke, M. S. Schmid, M. Keijzer and S. Dostert (Amsterdam: John Benjamins Publishing Company), 135–153.
- Schmid, M. S. (2002). First language attrition, use and maintenance. *The case of German Jews in Anglophone Countries. Studies in Bilingualism* 24. Amsterdam: John Benjamins.
- Schmid, M. S. (2011). *Language attrition*. Cambridge: Cambridge University Press.
- 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. (2017a). The relevance of first language attrition to theories of bilingual development. *Linguistic Approaches to Bilingualism*. 7, 637–667. doi: 10.1075/lab.17058.sch
- Schmid, M. S., and Köpke, B. (2017b). When is a bilingual an attriter? Linguistic approaches to development. *Linguistic Approaches to Bilingualism*. 7, 763–770. doi: 10.1075/lab.17059.sch
- Smeets, L. (2022). "Feature reconfiguration at the syntax–discourse interface: L2 acquisition of Italian CLLD" in *Generative SLA in the age of minimalism: Feature, interfaces, and beyond*. eds. T. Leal, E. Shimanskaya and C. Isabelli (Amsterdam: John Benjamins Publishing Company), 213–240. Available at: <https://journals.sagepub.com/doi/full/10.1177/02676583231156307>
- Smeets, L. (2023). Feature reassembly and L1 preemption: acquiring CLLD in L2 Italian and L2 Romanian. *Second. Lang. Res.* 02676583231156307. doi: 10.1177/02676583231156307
- Sorace, A. (1993). Incomplete vs. divergent representations of unaccusativity in native and non-native Italian grammars of Italian. *Second. Lang. Res.* 9, 22–47. doi: 10.1177/026765839300900102
- Sorace, A. (2000a). "Differential effects of attrition in the L1 syntax of L2 near-native speakers" in *In 24th annual Boston University conference on language development*. eds. C. Howell, S. Fish and T. Keith-Lucas (Somerville, MA: Cascadia Press), 719–725.
- Sorace, A. (2000b). Syntactic optionality in non-native grammars. *Second. Lang. Res.* 16, 93–102. doi: 10.1191/02676580067066032
- Sorace, A. (2003). "Near-nativeness" in *Handbook of second language acquisition*. eds. M. Long and C. Doughty (Oxford: Blackwell), 130–152.
- Sorace, A. (2011). Pinning down the concept of 'interface' in bilingualism. *Linguistic Approaches to Bilingualism*. 1, 1–33. doi: 10.1075/lab.1.1.01sor
- Stocchiero, A. (2002). Migration flows and small and medium sized Enterprise: Internationalization between Romania and the Italian Veneto region. Rome: CeSPI working paper.
- Tsimpli, I. M. (2007). "First language attrition from a minimalist perspective: Interface vulnerability and processing effects" in *Language attrition: A theoretical perspective*. eds. B. Köpke, M. S. Schmid, M. Keijzer and S. Dostert (Amsterdam: John Benjamins Publishing Company), 83–98.
- Tsimpli, I., Sorace, A., Heycock, C., and Filiaci, F. (2004). First language attrition and syntactic subjects: a study of Greek and Italian near-native speakers of English. *Int. J. Biling.* 8, 257–277. doi: 10.1177/13670069040080030601
- Villalba, X. (2000). *The syntax of sentence periphery* [unpublished doctoral dissertation] [Barcelona, (SP)]: Universitat Autònoma de Barcelona.
- White, L. (2023). "Assessing adult linguistic competence" in *The Routledge handbook of experimental linguistics*. eds. S. Zuffrey and P. Gyga (London: Routledge), 330–344.
- Wilson, F. (2009). *Processing at the syntax–discourse interface in second language acquisition* [Doctoral dissertation]. [Edinburgh (UK)]: University of Edinburgh.
- Wilson, F., Keller, F., and Sorace, A. (2009). "Antecedent preferences for anaphoric demonstratives in BUCLD 33" in *Proceedings of the 33rd annual Boston University Conference on Language Development*. eds. J. Chandlee, M. Franchini, S. Lord and G. Rheiner (Somerville, MA: Cascadia Press), 634, 645.
- Yılmaz, G., and Schmid, M. S. (2012). L1 accessibility among Turkish-Dutch bilinguals. *Mental Lexicon*. 7, 249–274. doi: 10.1075/ml.7.3.01yil



OPEN ACCESS

EDITED BY

Tania L. Leal,
University of Arizona, United States

REVIEWED BY

Radek Šimik,
Charles University, Czechia
Elena Shimanskaya,
University of Nevada, United States

*CORRESPONDENCE

Tatiana Luchkina
✉ tatiana.luchkina@stonybrook.edu

RECEIVED 31 December 2023

ACCEPTED 06 September 2024

PUBLISHED 24 December 2024

CITATION

Luchkina T, Ionin T and Goldshtein M (2024)
Acquisition of non-contrastive focus in
Russian by adult English-dominant bilinguals.
Front. Psychol. 15:1363980.
doi: 10.3389/fpsyg.2024.1363980

COPYRIGHT

© 2024 Luchkina, Ionin and Goldshtein. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Acquisition of non-contrastive focus in Russian by adult English-dominant bilinguals

Tatiana Luchkina^{1*}, Tania Ionin² and Maria Goldshtein³

¹Linguistics Department, Stony Brook University, Stony Brook, NY, United States, ²Linguistics Department, University of Illinois at Urbana-Champaign, Urbana, IL, United States, ³Ira A. Fulton Schools of Engineering, Arizona State University, Tempe, AZ, United States

This study investigates the acquisition of sentence focus in Russian by adult English-Russian bilinguals, while paying special attention to the relative contribution of constituent order and prosodic expression. It aims to understand how these factors influence perceived word-level prominence and focus assignment during listening.

We present results of two listening tasks designed to examine the influence of pitch cues and constituent order on perceived word prominence (Experiment 1) and focus assignment (Experiment 2) during the auditory comprehension of SV[O]_F and OV[S]_F sentences in Russian. Our findings reveal an asymmetric pattern: monolingual speakers, as a baseline, tend to perceive the nuclear pitch-accented object as more prominent, particularly in the SVO order, whereas bilinguals appear to be less sensitive to the constituent order distinction.

Additionally, baseline speakers consistently assign focus to the sentence-final nuclear pitch-accented noun regardless of constituent order. In contrast, bilinguals demonstrate a preference for assigning focus to the sentence-final nuclear-accented object, rather than the sentence-final nuclear-accented subject. A proficiency effect emerged indicative of a more target-like performance among bilinguals with greater proficiency in Russian.

KEYWORDS

focus, information structure, prosody, constituent order, Russian

1 Introduction

The present study critically evaluates the ability of adult English-Russian bilinguals to infer sentence focus in Russian, a free word order language, in both canonically ordered SVO sentences and non-canonically ordered OVS sentences, during auditory sentence comprehension.

Similar to other languages with pitch accents, Russian exhibits prosodic effects tied to the information status of referents, which is reflected in pitch accent patterns at the phrasal level. This includes emphasizing new, focused information while de-emphasizing given information (Neeleman and Titov, 2009; Jasinskaja, 2016).

The Nuclear Stress Rule (NSR; Chomsky and Halle, 1968) establishes that the main phrasal prominence, or nuclear pitch accent, is placed at the rightmost prosodic domain boundary. In both Russian and English, Intonational Phrases (IPs) define the prosodic domain within which the NSR operates. Russian shares similarities with English in that focusing a non-IP-final word shifts the nuclear pitch accent to a non-phrase-final position to align with the focused

constituent. However, unlike English, Russian uses overt case morphology and alters the order of sentence constituents to convey information status and relative prominence (Bailyn, 1995; Kallestinova, 2007; Slioussar, 2007).

Consider the example provided in (1); the subject question in (1a) can receive a response with non-canonical OVS order, shown in (1b) with the focused subject noun “lisa” (eng.: “fox.NOM”) positioned at the end of the sentence and aligned with the nuclear pitch accent sentence-finally. A similar pattern emerges in the object question in (2a), which can receive a response in the baseline SVO order, with a sentence-final nuclear pitch accent (2b)¹. Examples 3 and 4 illustrate the same question-answer pairs in English; since OVS order is not possible in English, both (3b) and (4b) have SVO order, but the nuclear pitch accent is aligned with either the subject or the object, depending on the question type.

Here and below, CAPS in example sentences indicate the nuclear accent and _F represents the focus.

- | | |
|---|--|
| 1a. Kto napugal volka?
who.Nom scared wolf.Acc
'Who scared the wolf?' | 2a. Kogo napugala lisa?
who.Acc scared fox.Nom
'Who did the fox scare?' |
| 1b. Volka napugala [LISA] _F .
wolf.Acc scared fox.Nom
'A/the fox scared the wolf' | 2b. Lisa napugala [VOLKA] _F .
fox.Nom scared wolf.Acc
'The fox scared a/the wolf' |
| 3a. Who scared the wolf?
3b. The FOX scared the wolf. | 4a. Who did the fox scare?
4b. The fox scared the WOLF. |

Example 1 demonstrates that in Russian, the focal reading of a sentence element allows for its relocation to the final position in the utterance which serves as the primary landing site for the main phrasal prominence, i.e., the nuclear pitch accent. While baseline monolingual speakers of Russian demonstrate sensitivity to focus-driven constituent order alternations (Laleko, 2022) and exhibit variability in prosodic expression linked to information status distinctions (Luchkina and Cole, 2021), heritage speakers and second language (L2) learners of Russian have been found to display non-native-like patterns of acceptability for non-canonical orders (Laleko, 2022; Ionin et al., 2023a; Ionin et al., 2023c).

This observation suggests that the concurrent use of constituent order and prosody in the expression of sentence focus may pose a challenge for adult L2 speakers of Russian. Acquisition challenges may stem from learners encountering difficulties in identifying sentence focus through prosodic cues, or in integrating word order and acoustic-prosodic expression with the discourse context (Ionin et al., 2023a).

The Interface Hypothesis for second language acquisition (Sorace and Filiaci, 2006) predicts increased complexity and resistance in acquiring properties that require the integration of language-internal

and language-external domains, including syntax and information structure.

Interface phenomena investigated by Sorace (2011), along with much subsequent research, exhibited optionality in the interlanguage (IL) grammars of adult L2 learners, which contrasts with the more consistent grammars of native (baseline) speakers. For instance, Sorace (2007) examined the use of overt and null subjects by highly proficient L1 English learners of L2 Italian. In Italian, the use of null subjects is determined by the information status of the subject (new vs. given), with discourse-given subjects typically being null. Sorace reported residual optionality (i.e., the use of both overt and null subjects) among the tested L2 learners in contexts where native speakers consistently opted for null subjects. Sorace attributed this optionality in the use of null subjects by L1 English learners of L2 Italian to the complex nature of the interface between syntax and discourse that is inherent in this aspect of Italian syntax.

In the context of the present study, information structure in L2 Russian requires the simultaneous use of distinct target language properties, including constituent order and prosodic expression. The associated acquisition challenges documented in earlier research may therefore stem from learners' difficulties in identifying sentence focus through prosodic cues, or from the integration of word order and acoustic-prosodic expression with the discourse context, as predicted by the Interface Hypothesis.

The present study assesses these possibilities by comparing monolingual speakers of English and Russian with adult English-Russian bilinguals whose dominant language is English. These speaker groups are compared on their perception of the main phrasal prominence in sentences such as (1b) and (2b) and further, on concurrent use of constituent order and prosodic expression as cues to focus assignment during listening.

2 Expression of sentence focus

2.1 Pitch accenting

In the influential research by Chafe (1976), focus is defined as an intrinsic attribute of the utterance information structure. In the present study, we use the term “focus” to signify newly introduced information within a sentence that is expected to be the primary point of interest for the listener or reader [see Cruschina (2022) for more discussion].

In pitch accenting languages, including Russian and English, focal information tends to be prosodically distinct due to relative prosodic augmentation of the sentence focus in combination with partial reduction of prominence of non-focal, given information. Extensive foundational research on spoken English has established a clear link between heightened information emphasis, often attributed to focal status, and prosodic prominence (Beckman and Pierrehumbert, 1986; Selkirk, 1995; Ladd, 2008; Büring, 2009; Wagner et al., 2010; Cole, 2015; Bishop et al., 2020).

Sentence focus frequently exhibits a distinct prosodic expression, thus rendering it prosodically prominent, as discussed in the works of Selkirk (1995), Ladd (2008), Büring (2009), Calhoun (2010), and Bishop et al. (2020). In English, focus prominence results from distinctive pitch accenting patterns linked to the relative information prominence of a

¹ The question in (1a) can also receive a response in SVO order, in which case the nuclear pitch accent would be on the preverbal subject; in principle, (2a) can receive a response in OVS order, with the nuclear pitch accent on the preverbal object, though this is rather less commonly attested [see Kallestinova (2007)]. We leave this aside, since the studies reported in this paper examined exclusively sentences with nuclear pitch accent on the sentence-final constituent, as in (1b) and (2b).

word. When a word holds focal status, it is assigned a nuclear pitch accent, effectively linked to the most perceptually salient prosodic event within a larger domain, such as an IP. The form of the pitch contour indicating focus or discourse-new information status is informed by the specific pitch accent type, such as H* (Pierrehumbert and Hirschberg, 1990; Beckman et al., 2005).²

Perception-production studies by Gussenhoven and Rietveld (1988), Xu and Xu (2005), Breen et al. (2010), and Bishop et al. (2020) reported significant contribution of the local pitch maxima, the speed of pitch rise and the size of pitch excursion over the focused word to acoustic-prosodic expression of focus in English. In perceptual terms, the augmented prosodic expression translates into heightened prosodic prominence of the focal material (Xu and Xu, 2005; Cole, 2015), which may further translate into variable degrees of perceived information prominence by linguistically naïve listeners (Breen et al., 2010).

Given the various mechanisms that contribute to the focus prominence, studies examining the production and perception of sentence focus reveal inherent variability in how speakers express it orally, as well as in how listeners perceive it (Breen et al., 2010; Takahashi et al., 2018).

The work of Breen et al. (2010) presents an illustrative perception-production study of English focus. In their comprehensive analysis of the acoustic-prosodic focus correlates, the authors emphasized the crucial role of several acoustic parameters, including pitch, loudness, and segmental length in distinguishing the focused element from the rest of the sentence. In a series of discriminant function focus identification analyses, prosodic expression helped determine the location of the focused word in test sentences but proved insufficient to discriminate between contrastive (LH*) vs. non-contrastive focus (H*) or determine the size of the focus domain (broad vs. narrow). Linguistically naïve listeners tested by Breen et al. (2010) were highly successful at locating the sentence focus (10/10 succeeded) but only moderately successful at identifying the focus type (contrastive vs. non-contrastive, 6/10 succeeded) or the focus domain size³ (8/10 succeeded).

The same study by Breen and colleagues analyzed read production-perception data from 13 unique sets of speakers. In each speaker pair, partner 1 read a target sentence and partner 2 selected one of the seven questions for which participant 1's production served as the most plausible answer.⁴ The authors reported an overall accuracy of 55%, which was above chance given the large number of context options available to the listeners. About half (46%) of Breen et al.'s (2010) participants achieved above chance accuracy at identifying wide focus, and 70% of the participants were above chance at identifying narrow object focus phrase-finally.

The prosodic correlates of sentence focus in Russian have been investigated by Bryzgunova (1980) and Zybatow and Mehlhorn

(2000). In a more recent review by Jasinskaja (2016), the prosodic analysis of Russian focus is grounded in a detailed examination of intonational patterns, pitch accents, and their interaction with syntactic structure and discourse context. Jasinskaja (2016) bases her prosodic analysis of Russian focus on Bryzgunova's (1980) pitch (intonational) contour classification, originally developed for categorizing "neutral" and "non-neutral" intonation patterns in Russian. Using Bryzgunova's terminology, the *neutral intonational contour* pertains to the SVO sentence pattern with new information focus positioned toward the end of the phrase or aligned with its rightmost edge. In terms of prosody, sentences aligns clause-final new information focus can feature several down-stepped pre-nuclear pitch accents on each pre-focal word. The H tone of the HL* bitonal nuclear accent aligns with the pretonic syllable of the focused word exponent, leading to a drop in pitch over the stressed syllable.

Word-level augmented prosodic expression of non-contrastive new information foci was identified as a reliable predictor or perceived information prominence by adult Russian listeners by Luchkina and Cole (2021). This effect was further amplified by variations in word order, a topic we explore next.

2.2 Word order

Due to the relatively free constituent order in the Russian language, Kallestinova (2007) and more recent studies by Luchkina and Cole (2021), Ionin et al. (2023a), and Laleko (2022, 2024) have investigated the role of word order in signaling sentence focus. This research has demonstrated that the SVO and OVS orders in Russian correspond to distinct configurations in terms of information structure. In the baseline SVO order, the subject is generally assumed to be part of the ongoing discourse, while the object is considered new information and is in focus. Conversely, with OVS order, the object is established in the discourse, i.e., is topical, while the subject takes center stage in the listener's attention, i.e., is in focus.

SVO [as in (2b)], which is typically seen as the default word order (Bivon, 1971), can be adjusted prosodically to suit different IS scenarios (Laleko, 2024). In contrast, OVS [as in (1b)] appears more marked (Sekerina, 1999) and necessitates an interpretative license (Titov, 2017). According to Kallestinova's experimental research on constituent orders in Russian, speakers use OVS when they want to emphasize the subject, but not in other contexts (Kallestinova, 2007). This suggests that the limited applicability of non-standard word orders renders them less amenable to prosodic adjustments Luchkina et al., (in prep). Overall, the experimental evidence points to a distinct interplay between word order and prosody in Russian, particularly in scenarios involving subject and object focus.

Figure 1 offers illustrative pitch contours associated with baseline SVO order in and the subject-final OVS order. Both pitch tracks illustrate clause-final placement of the nuclear pitch prominence.

3 L2 acquisition of sentence focus

While constituent order and prosodic cues are generally reliable indicators of distinctions in information structure for native speakers, they pose a recognized difficulty in the acquisition for adult L2 learners and heritage language speakers.

² In line with Calhoun (2010), Katz and Selkirk (2011) treat both H* and LH* as accents marking new information focus in English.

³ The focus distinctions in question arise due to distinctions in focus breadth. Broad focus renders two or more adjacent words focal in a phrase or an utterance. Narrow focus assumes the focal status of a single word in a sentence or phrase.

⁴ Breen et al. (2010) tested various focus conditions, including broad vs. narrow focus, phrase-final vs. phrase-medial narrow focus, and focus type (contrastive vs. non-contrastive). Only the results relevant for the present study are summarized.

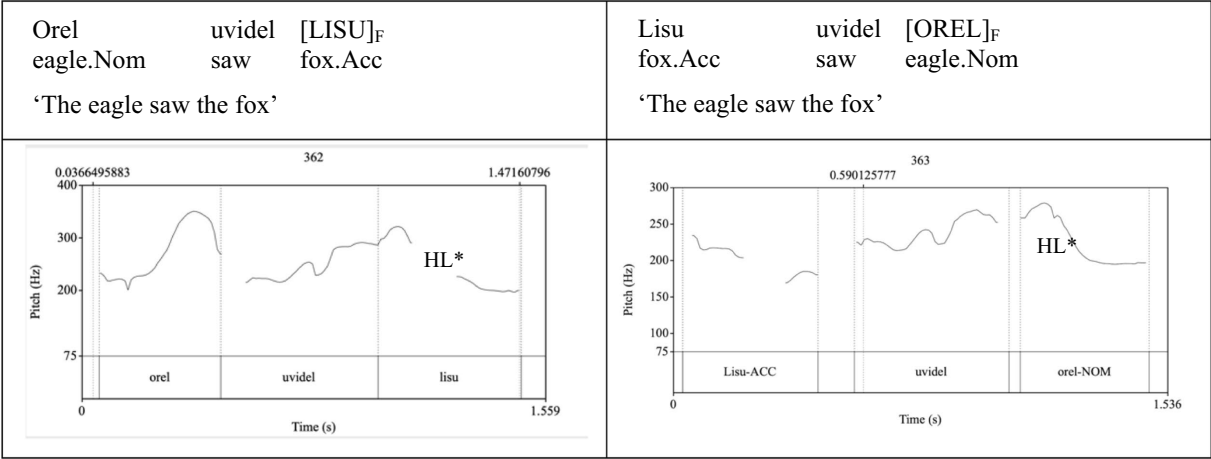


FIGURE 1
Illustrative pitch tracks of an SVO sentence (left) and an OVS sentence (right) ‘The eagle saw the fox’ with HL* clause-final nuclear pitch.

3.1 Known acquisition challenges

One potential source of difficulty in identifying focus based on auditory cues is the subtle and variable nature of the nuclear pitch prominence, which serves as the acoustic-prosodic expression of new information focus. Although nuclear pitch accent is often cited as the most reliable cue to sentence-level prominence and focus (e.g., Gussenhoven, 2004), no single acoustic correlate of nuclear pitch prominence has been established for English, Spanish, or Russian, such that it would enable identification of the prominent word directly from quantitative acoustic measurements, without an auditory analysis (Beckman, 1996; Fletcher and Evans, 2002).

The probabilistic nature of focus expression during speech underlies a great deal of individual variability in focus perception and production. A recent study by Takahashi et al. (2018) compared native English speakers and L1 Mandarin L2 English learners on the production and perception of narrow contrastive focus in English. When examining the use of acoustic-prosodic expression in relation to the focused constituent, it was observed that not all baseline speakers chose to produce a nuclear pitch accent in the vicinity of the sentence focus. Furthermore, expressing focus through prosody during production did not determine whether the same speaker relied on acoustic-prosodic cues for auditory focus identification. Takahashi et al. reported a similar ‘disconnect’ between the production and perception of English contrastive focus in a group of proficient L1 Mandarin L2 English speakers. These findings highlight the inherent inter-speaker variability and the probabilistic nature of the prosodic cues used to convey focus in the context of L1-L2 English.

Another challenge may arise from the relatively greater complexity of focus expression, which involves bridging multiple domains, including prosody, syntax, and information structure.

The Interface Hypothesis (IH), formulated by Sorace and Filiaci (2006), provides a theoretical framework that underscores the challenges associated with acquiring phenomena at the intersection of a language-internal syntax domain and language-external phenomena, including information structure. Due to the greater underlying complexity of interface phenomena, the targeted interface structures

resist acquisition, even in IL grammars of learners at advanced proficiency levels. In particular, Sorace (2011) proposed that the acquisition of external interface phenomena is linked to instances of optionality within the target grammar, as well as “protracted indeterminacy” found even in near-native L2 learners (2011, p. 5). This stands in contrast to structures that are purely language-internal and, therefore, more readily acquirable.

Experimental evidence supporting the IH emphasizes the transfer of focus marking strategies from the native language, as shown in studies by Hertel (2003), Fruit (2006), and Ortega-Llebaria and Colantoni (2014). To illustrate, Fruit (2006) examined how Brazilian Portuguese (BP) speakers of varying English proficiency levels interpreted different focus structures in both BP and English. Fruit observed that even the L2 speakers considered to have achieved near-native proficiency in the TL exhibited L1 influence in their selection of constituent order and accent placement. For example, some participants showed a preference for sentence inversion in cases of narrow focus, which deviated from the standard SVO order used in combination with prosodic emphasis on the focused word in the TL. Fruit concluded that the interface between syntax and information structure presents a challenge in acquisition, even for L2 learners whose TL syntax is generally similar to the native language. Fruit identified L1-biased optionality and transfer from L1 to L2 as probable factors contributing to the variable performance observed among the tested participants, even among those who otherwise exhibited convergence with the TL syntax.

In a similar vein, Ortega-Llebaria and Colantoni (2014) found that learners’ focus marking patterns in their native language may result in lasting transfer effects when acquiring an L2, regardless of proficiency level. The authors examined how native Spanish and native Mandarin speakers learning English as an L2 perceived sentence focus. Both groups were tasked with identifying the location of the word in contrastive focus in sentences presented with or without context. The study found that native Mandarin speakers demonstrated a high level of accuracy, closely resembling native English speakers, even though their overall proficiency in English was lower than that of native Spanish speakers. This accuracy likely stemmed from positive transfer from their native

language, which, like English, employs prosodic prominence to convey contrastive focus. In contrast, native Spanish speakers, despite their higher TL proficiency, showed noticeably lower accuracy. This discrepancy may be attributed to the Spanish tendency to use word order to position the focused element at the end of the sentence, where it receives the nuclear pitch accent. As expected, the accuracy of L1 Spanish speakers was greater when the focused element was the post-verbal object compared to when it was the pre-verbal subject, underscoring the strong influence of the L1 focus-marking strategies.

Notably, several experimental investigations of production and perception of sentence focus brought forward evidence supporting that L2 focus is acquirable and that successful acquisition critically depends on the proficiency in the TL. To illustrate, a production study on inverted VS order in the expression of Spanish information structure conducted by Hertel (2003) with L1 English L2 Spanish learners revealed a presence of L1 transfer from English, particularly at lower and intermediate levels of TL proficiency. An emerging sensitivity to discourse factors, including focus, was observed in advanced-level learners who demonstrated a native-like preference for the VS constituent order used to signal subject focus.

In summary, challenges in acquiring L2 focus may stem from differences in linguistic means used as focus cues between the speakers' L1 and the TL (e.g., information structure primarily interfaces with constituent order in Spanish but with phrasal prosody in English). As predicted by the Interface Hypothesis, when the domain in question—focus—intersects language-internal and language-external elements, that domain becomes vulnerable and resistant to acquisition. This intersection complicates the learning process, making it more challenging to fully acquire the relevant interface structures.

3.2 Evidence from L2 Russian

In recent years, several experimental investigations have focused on the acquisition of the information structure and its effects on constituent order in Russian. Ionin et al. (2023a) and Laleko (2022) both conducted acceptability judgment studies, where English-dominant English-Russian bilinguals evaluated the acceptability of baseline SVO and inverted OVS stimuli sentences, considering the focal reading of one of the nominal constituents.

Ionin et al. (2023a) utilized pre-recorded auditory SV[O]_F and OV[S]_F test sentences featuring narrow focus clause-finally. Acceptability patterns varied among bilinguals, with heritage speakers, but not adult L2 learners, interpreting the OVS order as a means of signaling subject focus. Laleko (2022) reported similar results using written stimuli sentences. Laleko's study, similarly, found that heritage bilingual speakers with higher Russian proficiency, but not adult Russian L2ers, succeeded at accepting the subject-final order in transitive OV[S]_F sentences with subject focus.

Laleko (2022) extended her investigation into the information structure domain in heritage Russian by assessing the acceptability of pre-recorded SVO and OVS sentences. Focus in these sentences was marked either through prosodic cues or constituent reordering. An asymmetry surfaced, where baseline monolingual speakers exhibited no preference for either focus marking strategy. In

contrast, heritage speakers clearly favored nuclear pitch accenting of the focused word *in situ*, rendering constituent reordering redundant. The same study reported that heritage speakers over-accepted phrase-final nuclear pitch prominence under narrow subject focus in the SVO order. In contrast, under object focus, they correctly rejected infelicitous placement of the nuclear accent in the sentence-initial position. Laleko (2022) interpreted these findings as evidence of partial "neutralization in prosodic patterns" by heritage Russian speakers (p. 16).

Recent evidence supporting on-target perception of prosodic cues in relation to contrastive sentence focus in Russian was presented by Ionin et al. (2023b) who tested contrastive focus (CF) identification in SV[O]_{CF}, S[O]_{CF}V, and [S]_{CF}VO experimental sentences, preceded by a one-sentence discourse context. The study found that adult L2 learners of Russian successfully identified the word in contrastive focus, regardless of whether it occurred sentence-finally (SV[O]_{CF}) or elsewhere (e.g., S[O]_{CF}V).

Experiment 2 in Ionin et al. (2023b) assessed focus identification during silent reading, requiring listeners to rely solely on context cues, and during listening, where the word in focus was made prosodically prominent. In the listening phase, both felicitous and non-felicitous contexts were examined to measure listeners' ability to determine the location of nuclear pitch prominence in the absence of supportive context cues. The study reported a notably accurate performance from 26 adult English-Russian bilinguals, with above 90% accuracy during silent reading and listening. During listening, identification accuracy remained well above chance even when the target sentences were presented along with non-felicitous contexts (the context sentence set a non-nuclear accented noun in focus). This reveals listeners' sensitivity to the prosodic cues in expression of contrastive focus under various constituent orders and phrasal locations. Ionin et al. (2023b) reported that the participants' TL proficiency served as a crucial predictor of accurate contrastive focus identification during listening.

Contrary to Ionin et al. (2023b), a related investigation of non-contrastive focus in Russian, by Luchkina et al. (in press), reported considerably more indeterminacy on part of both native Russian listeners and adult English-Russian bilinguals when these groups were tested on auditory comprehension of SV[O]_F and OV[S]_F sentences featuring an instance of non-contrastive new information focus clause-finally. Near-ceiling accuracy was achieved during the silent reading task. During listening, participants had to identify the most prosodically prominent word in the test sentences presented with context. The study reported a 57% rate of perceived nuclear prominence in felicitous question-answer pairs (object question followed by an SV[O]_F target) and a low 28% accuracy on non-felicitous question-answer pairs (object question followed by an OV[S]_F target) in their data from English-Russian bilinguals. These findings reveal that bilinguals were more likely to identify the nuclear-accented word as prominent when it was in focus. A follow-up analysis, in which participants' TL proficiency was considered, revealed robust prominence identification in congruent question-answer pairs but a proficiency-dependent outcome for non-congruent, illicit question-answer pairings. The context felicity effect and its interaction with TL proficiency jointly point to less proficient bilingual listeners relying on context cues, rather than prosodic cues, for focus identification.

In summary, prior experimental investigations have emphasized the complexity of acquiring sentence focus in Russian as the TL. This complexity arises from the involvement, on one hand, of distinct language-internal means of signaling focus in spoken language use, including prosodic cues and constituent order, and on the other, coordinating these mechanisms with the language-external domain of information structure.

4 The present study

In this study, we further explore the perception of nuclear pitch prominence using the test stimuli from [Ionin et al. \(2023a\)](#) and [Luchkina et al. \(in press\)](#). We present novel perception data from monolingual reference groups of Russian and English listeners, and English-dominant bilinguals. Considering the TL proficiency effect on the perceptual judgments of sentence prominence reported in the earlier related work, this investigation focuses on English-Russian bilinguals whose scores from an independent proficiency measure (cloze deletion test) substantiate a significant level of proficiency in Russian.

The present study is guided by two primary objectives. The first objective is to investigate perceived word-level prominence based on (1) tonal cues associated with the auditory expression of sentence focus in the languages spoken by the population of interest: English (dominant) and Russian (non-dominant) bilinguals and (2) linear order of the sentence constituents in Russian. The second objective of this study is to empirically evaluate how both constituent order and tonal cues linked to nuclear pitch prominence are employed concurrently in Russian during focus assignment, a task undertaken by native speakers and English-Russian bilinguals.

We begin, in Experiment 1, by assessing the perceptual weight of auditory cues to phrase-final nuclear pitch prominence across three speaker groups—two monolingual reference groups of English and Russian speakers, and a test group consisting of English-Russian bilinguals. This analysis aids in assessing whether the non-target-like performance exhibited by bilingual speakers, as noted in previous studies ([Ionin et al., 2023a](#); [Luchkina et al. in press](#)), can be linked to the perception of tonal cues to word-level prosodic prominence in Russian by English-dominant bilinguals.

To this end, in Experiment 1, we assess whether proficient English-Russian bilinguals demonstrate a target-like use of prosodic cues to phrase-final nuclear pitch prominence in Russian SVO and OVS stimulus sentences. The following research questions are addressed: (1) What cues underly the percept of prosodic prominence in simple transitive Russian sentences? (2) Do English-Russian bilinguals align with baseline Russian speakers in their prominence ratings?

The second objective of the present study is to empirically assess the concurrent use of constituent order and tonal correlates of the nuclear pitch prominence during focus assignment in Russian. This is achieved by testing focus assignment preferences in the reference group of Russian monolinguals and the test group of English-Russian bilinguals.

To this end, in Experiment 2, we carry out a backward focus assignment task (originally implemented with English NSs in [Green et al., 2010](#)). In this task, listeners use prosodic expression and constituent order in spoken test sentences as heuristics for detecting which word presents the most likely non-contrastive focus exponent in the given sentence. The following research questions are addressed: (3) Which cues do listeners rely on during focus assignment? (4) Do English-Russian bilinguals align with baseline Russian speakers in their use of prosodic expression and constituent order when assigning focus at phrasal level?

4.1 Experiment 1: the prominence identification task

Experiment 1 tested perception of phrasal prominence in simple transitive sentences presented without supporting context. Given that English-Russian bilinguals have previously demonstrated indeterminacy in auditory prominence identification ([Luchkina et al., in press](#)) and non-target-like, limited acceptance of focus configurations under the OVS constituent order ([Ionin et al., 2023a](#)), we compare word-level prominence ratings from monolingual speakers of Russian and English, recruited as reference raters, to those from a group of English-Russian bilingual listeners. Critically, the present study purposefully zooms in on the contribution of the prosodic expression, with a special focus on the tonal cues (per prior account of the Russian intonation by [Bryzgunova, 1980](#)), to perceived prosodic prominence at word level. For that reason, the stimuli sentences in Exp. 1 are presented for prosodic prominence identification as stand-alone, no context provided, utterance-long segments.

4.1.1 Materials

The Russian stimuli sentences comprised 24 pre-recorded SVO sentences (e.g., 1b. repeated as 5a. below) and 24 pre-recorded OVS sentences (e.g., 2b. repeated as 5b. below) from [Ionin et al. \(2023a\)](#) and [Luchkina et al. \(in press\)](#). The nuclear pitch prominence in the pre-recorded test sentences (HL*) was invariably phrase-final, i.e., aligned with the object nominal in the SVO order and subject nominal in the OVS order. All subjects and objects were animate nouns. All objects contained an overt Accusative case marker. For the English version of the task, the Russian SVO test sentences ($n = 24$) were translated into English and audio recorded (see example 3 repeated in 6. below).

- | | |
|--|--|
| 5a. Volka napugala [LISA] _F .
wolf.Acc scared fox.Nom
'A/the fox scared the wolf' | 5b. Lisa napugala [VOLKA] _F .
fox.Nom scared wolf.Acc
'The fox scared a/the wolf' |
|--|--|

6. The fox scared the WOLF.

Both Russian and English item lists included fillers originally recorded and tested in [Ionin et al. \(2023b\)](#). Each filler was one sentence long and contained a contrastively accented word in variable phrasal positions (LH* in Russian; L + H* in English). In the Russian version of the task, fillers featured variable constituent orders. These included SV[O]_F and S[O]_FV, [S]_FVO

and S[V]_FO configurations (see examples 7a–d). English fillers were SVO sentences with variable placement of the nuclear pitch accent:

[S]_FVO, S[V]_FO, SV[O]_F.

7.

- | | | | | |
|----|----------------------|-----------------------|-----------------------|-----------------------|
| a. | Vasya | poshel | v | [SHKOLU] _F |
| | Vasya.NOM | went | to | school |
| b. | Vasya | v | [SHKOLU] _F | poshel |
| | Vasya.NOM | to | school | went |
| c. | [VASYA] _F | v | shkolu | poshel |
| | Vasya.NOM | to | school | went |
| d. | Vasya | [POSHEL] _F | v | shkolu |
| | Vasya.NOM | went | to | school |
- ‘Vasya went to school.’

During stimuli recording sessions, on the speakers’ reading sheet, each target sentence was preceded by one-question-long context (see examples 1–4) which set the sentence-final noun in focus. The model speakers were instructed to read the question-answer pairs with natural intonation, with main prominence on the sentence-final noun. Only the answer component of each item (target or filler) was utilized in the listening tasks reported in this study.

The model speakers were female native speakers of Russian and English who did not participate in any of the tasks. The English speaker was not informed about the purpose of this study and was not linguistically trained. The Russian speaker served as an investigator on an earlier project involving the same set of stimuli (Ionin et al., 2023a) and was a graduate student in Linguistics when recordings were made.

The model speakers read the target sentences with neutral intonation, with main prominence on the sentence-final noun. For filler items, recorded subsequently, the location of the main phrasal prominence was indicated using UPPER CASE letters on the speaker’s reading sheet. Recordings were completed in a soundproof booth, at the University of Illinois Phonetics lab.

The 24 English target sentences were presented in a single item list, intermixed with 24 English filler sentences. There were two item lists in the Russian task each containing twenty-four fillers and 24 target sentences, 12 SVOs and 12 OVSs.

4.1.1.1 Acoustic-prosodic analyses of the recorded stimuli sentences

The recorded audio was digitized at a sampling rate of 44 k, and manually annotated in Praat (Boersma and Weenink, 2024). Several tonal correlates of the main phrasal prominence were examined, including word-level pitch minima, maxima (Hz, st), and excursion size⁵ (st), extracted from each nominal constituent (subjects and objects).

All measures of interest were sampled twice. The first set of measurements was extracted from the stressed vowel in each content word in the experimental sentences. The second set comprised word-level measurements, which were not limited to the tonic vowel. For the inferential analyses in the present study, we opted to use word-level measures in set 2. This decision was based on earlier work on Russian, which found that the post-tonic syllable often aligns with a pitch peak or another critical element of the pitch contour [see Jasinskaja (2016), for further discussion].

In the Russian stimuli sentences (see Table 1), object pitch maxima exceeded those of subjects in the object-first OVS order but not in the subject-first SVO order. Similarly, pitch excursion over the objects was greater than those over the subjects in the OVS order, but not in the SVO order.⁶ While none of the examined acoustic-prosodic parameters conclusively demonstrated quantifiable evidence of prosodic augmentation in the vicinity of the phrase-final nuclear pitch-accented noun,⁷ visual inspection of the pitch contours over the sentence-final noun revealed consistency with the HL* intonational contour, in line with the analysis of the Russian intonation by Bryzgunova (1980), Zybatow and Mehlhorn (2000), see Šimík (in press) for more extensive discussion.

The acoustic-prosodic measures extracted from the English stimuli (see Table 2) paralleled those reported for Russian, but also included an additional set of measures sampled from the sentence-medial verbal constituent. This was deemed necessary because in the English version of the prominence identification task, the verb was often rated as prosodically prominent.

Analyses of the acoustic-prosodic expression in the English stimuli revealed that verbs had, on average, the highest pitch peak values [max f0 = 72st (SD = 27.8st)], surpassing the highest pitch values over sentence subjects, on average, by 9st, and over objects, on average, by 21st. Meanwhile, sentence-final objects exhibited relatively lower pitch minima and maxima but displayed the greatest pitch excursion [mean = 17.1 (SD = 8.4 st)]. Where possible, visually inspected pitch contours over the sentence-final noun were consistent with the H* intonational contour, supporting prior analyses of English intonation (Katz and Selkirk, 2011).

Visual examination of the nuclear pitch peaks was not possible in all English stimuli sentences due to a high incidence of vocal fry in the

6 Gussenhoven and Rietveld (1988) point out that the magnitude of pitch excursion due to accenting is affected by phrasal position of the accented words. As a result, accents occurring later in the phrase exhibit smaller excursions than those occurring earlier.

7 For ease of comparison and in line with prior literature on Russian, our analyses focus exclusively on known tonal correlates of nuclear pitch prominence when modeling the experimental data. Two non-tonal acoustic measures—vocalic segment duration and intensity—were analyzed in the Russian production data. Previous research on Russian [see Luchkina and Cole (2021)] has shown that lower vowel intensity combined with greater segment duration often occurs in the vicinity of the nuclear accented word, likely due to its proximity to the phrase-final boundary. However, in our inferential analyses of perceived word-level prominence and focus assignment, neither intensity nor duration achieved statistical significance. Because a detailed interpretation of these results is beyond the scope of the present study, these measures were excluded from the statistical models.

5 Pitch excursion, measured in semitones, quantifies the degree of nuclear rise in the vicinity of the accented word.

TABLE 1 Summary statistics of model speaker’s production data by constituent order and prosodic parameter; the Russian stimuli sentences.

Constituent order	measure	mean, Hz (st)	SD, Hz (st)
OVS	max f0, object nominal	362.3 (61.8)	56.72 (14.4)
	max f0, subject nominal	308.14 (50.19)	71.8 (17.44)
SVO	max f0, object nominal	307.47 (53.2)	63.9 (11.05)
	max f0, subject nominal	396.36 (68.5)	77.77 (13.5)
OVS	min f0, object nominal	191.01 (32.1)	9.45 (5.8)
	min f0, subject nominal	186.4 (30.4)	10.56 (7.83)
SVO	min f0, object nominal	182.2 (31.5)	9.11 (1.58)
	min f0, subject nominal	202 35.1	16.1 (2.8)
OVS	pitch excursion, object nominal (st)	10.88	2.98
	pitch excursion, subject nominal (st)	8.2	3.0
SVO	pitch excursion (st), object nominal	8.78	2.73
	pitch excursion (st), subject nominal	11.4	3.16

vicinity of the phrase-final nuclear pitch-accented noun. A similar incidence of vocal fry in recorded English sentence stimuli has been reported by [Yeung et al. \(2019\)](#). Following [Wolk et al. \(2012\)](#), [Yeung et al. \(2019\)](#) discuss the intonational pattern, whereby the utterance-final nuclear accented noun exhibits pitch declination in combination with significant vocal fry, typical for expressing new information focus by young speakers of American English. The analysis of the English stimuli, recorded by a young female speaker of American English, aligns with this observation, despite the fact that our English model speaker did exhibit the more conventional H* contour, as depicted in [Figure 2](#). We infer that in the English stimuli, the presence of vocal fry, coupled with pitch lowering, accounts for the relatively lower pitch peak values in the nuclear-accented nouns.

In summary, the analysis of model speakers’ performance data revealed a distinctive tonal quality in nuclear pitch-accented words, primarily attributable to pitch contours, rather than peak height or excursion height. This distinction was observed when comparing tonal measures of the nuclear-accented sentence-final noun with those of the sentence-initial noun. Notably, none of the investigated tonal parameters provided conclusive evidence of prosodic enhancement in the vicinity of the phrase-final nuclear pitch-accented noun.

These findings suggest that perceptual outcomes in the prominence identification task are likely to vary. Some listeners may expect to locate the nuclear-accented word phrase-finally, while others may seek a recognizable pitch contour or acoustic-prosodic expression

at the word level to identify the main phrasal prominence. These variable expectations are reflected in the testable predictions outlined for Experiment 1 below.

4.1.2 Participants

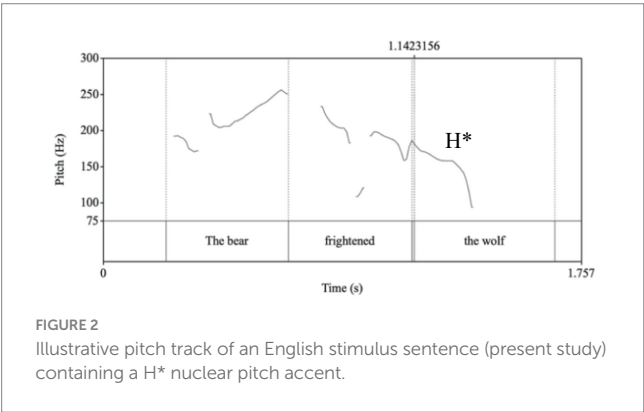
Data were obtained from three groups of linguistically naïve speakers, including baseline participant groups of Russian-speaking monolinguals (*n* = 29, mean age = 20.4) and English-speaking monolinguals (*n* = 68, mean age: 20.8). The monolingual speakers were recruited from among college student populations in Russia and in the US, respectively, and participated for course credit.

The third participant group included 29 English-Russian bilinguals (mean age = 36). The average age of exposure to English was 2.0 y.o; the age of exposure to Russian ranged between 0 and 30. All participants resided in the US, Canada, or Great Britain at the time of testing and declared English to be their native language as well as their preferred language for daily communication. Fifteen participants reported limited exposure to Russian via one or both parents. One participant reported that Russian was their native language, whereas English was their primary language. Thirteen participants reported completion of at least 2 semesters of formal classroom instruction in Russian as a foreign or heritage language. Seven additional bilingual participants were tested but eventually excluded due to extensive residence in a Russian speaking country (1 participant), failure to understand the task instructions (1 participant) and failure to meet the minimum

TABLE 2 Summary statistics of model speaker’s production data by constituent order and prosodic parameter; the English stimuli sentences¹.

Constituent order	Measure	Mean, Hz (st)	SD, Hz (st)
SVO	Max f0 (Hz), Object nominal	294.02 (50.9)	161.68 (27.85)
	Max f0 (Hz), Subject nominal	365.34 (63.25)	117.29 (20.3)
	Max f0 (Hz), Verb	417.63 (72.3)	168.67 (29.2)
	Min f0 (Hz), Object nominal	97.67 (16.1)	15.92 (2.76)
	Min f0 (Hz), Subject nominal	192.18 (33.27)	18.89 (3.27)
	Min f0 (Hz), Verb	184.3 (31.91)	40.06 (6.94)
	f0 excursion (st), Object nominal	17.082	8.44
	f0 excursion (st), Subject nominal	9.7	3.82
	f0 excursion (st), Verb	13.03	8.05

¹A referee points out greater distribution in the tonal values in the recorded English stimuli, in comparison to the Russian stimuli sentences. The difference, in perceptual terms, indicates a livelier reading style of the English model speaker. The Russian speaker’s narrow pitch range and overall smaller pitch excursions represent the speaker’s understanding of “neutral intonation”.



proficiency requirements (5 participants). The English-Russian bilinguals were paid for their participation.

4.1.3 Target language proficiency measures

The Russian language proficiency requirements for participant inclusion were established to guarantee comparable individual performance and avoid outcomes influenced by a deficit in TL proficiency, as previously noted by Luchkina et al. (in press) and in studies investigating focus perception in other languages (refer to Hoot, 2017 for relevant discussion). Furthermore, because the Russian stimuli involved a non-canonical constituent order, meaningful results can only be assured if the English-Russian bilinguals demonstrate above chance accuracy in interpreting the OVS stimuli as object-initial and subject-final, and not vice versa.

Bilingual speakers’ TL proficiency was evaluated using two independent measures, a 10-item test of morphological case previously

implemented in Ionin and Luchkina (2019), Ionin et al. (2023a) and a 57-item cloze deletion test (Luchkina et al., 2021). The correlation between these two proficiency measures in the present study (Pearson’s *r*) reached 0.69 (*p* < 0.0001). The case check test assessed participants’ accuracy in discriminating between the nominative and accusative cases based solely on overt morphological markers. The cutoff score for the case test was set at 0.6, and for the cloze deletion test - at 0.7. The mean accuracy achieved on the case check was 0.87. (range: 0.6–1. SD = 0.16); the mean accuracy achieved on the cloze deletion test was 0.77 (range: 0.71–0.96, SD = 0.15).

4.1.4 Procedure

All participants provided a written consent to participate and completed a language background questionnaire.

Participants were instructed to attentively listen to each target sentence and pay close attention to the prosodic expression in the model speaker’s read performance. Subsequently, participants were asked to select the word, in each target sentence, which they perceived as the most prominent, by clicking on it within the written sentence presented along with the audio recording.⁸

8 An anonymous reviewer points out the metalinguistic nature of the prominence identification task and the following focus assignment task (exp. 2) rendering this aspect of the methodology a limitation of the present study. The authors acknowledge the explicit nature of the tasks, deemed necessary to (1) determine the linguistic underpinnings of word-level prominence and (2) the link between perceived word-level prosodic expression and sentence focus. From the methodological standpoint, prosodic analysis by linguistically naïve listeners is used to draw inferences about the underlying prosodic grammar in the so-called rapid prosody transcription tasks (Cole and

In the prominence identification task, each content word could be selected as prominent, for each test sentence. Prominence, in this context, was defined as a word-level attribute that directs the respondent's attention more toward the prominent word compared to other words within the same segment. Instructions were presented in the dominant language of the task participants. Drawing on Cole et al. (2019, p. 120), in the English version of the task, prosodic prominence was characterized "as a word-level property leading certain words to have increased loudness, duration, pitch extremity, and 'crisper' articulation than the surrounding words." The Russian monolinguals were provided the following adaptation of the Cole et al.'s (2019) definition of prominence: "[...] select the word which the speaker highlighted by means of intonation. Such words are usually pronounced louder, longer, and with special voice timbre and may be regarded as key words in an utterance or phrase." Participants viewed two example items and completed three practice items, with feedback, before beginning the task. Participants completed this and the following tasks using Qualtrics online data collection platform.

4.1.5 Testable predictions

We predicted that all participants would opt for the word with the most prominent prosodic expression, attributed to pitch accenting or another salient prosodic property. Given the nuanced nature of nuclear pitch prominence in the phrase-final position (e.g., Katz and Selkirk, 2011), this might result in varying perceptual preferences. For instance, a non-phrase-final word could be perceived as prominent and not the nuclear pitch accented nominal. This potential outcome would be substantiated by relatively higher values of the tonal measures extracted from non-phrase-final constituents in the test sentences, as discussed in 4.1.1.1.

We anticipated that participant performance may differ based on the dominant language. Specifically, the Russian monolinguals were predicted to demonstrate sensitivity to the acoustic-prosodic expression, at word level, in the recorded test sentences. Furthermore, because two types of nominal constituents, subjects and objects, aligned with the phrase-final, nuclear pitch-accented position, we also predicted that listeners' judgements may be affected by constituent order in the test sentences as previously shown in Luchkina et al. (2015) and Luchkina and Cole (2021) who demonstrated that *ex-situ* words had a greater likelihood of being perceived as prominent by native Russian listeners. This suggests, for the OVS test sentences, a possibility for not just the nuclear accented subject, but also the fronted object, to be perceived as audibly prominent.

The English-Russian bilinguals were predicted to demonstrate sensitivity to the acoustic-prosodic expression at word level, due to transfer from the dominant language, more than to the constituent order when selecting the prominent word. This expectation arose from the lack of the OVS order in English.

Shattuck-Hufnagel, 2016; Cole et al., 2019; Bishop et al., 2020, among many others). The backward focus assignment tested in Exp. 2 presents a simplified version of a focus assignment task originally developed for English by Breen et al. (2010). Arguably, the methods chosen for the present study present straightforward, replicable means of testing auditory perception as well as integration of phrasal prosody and constituent order with the extra-linguistic domain of information structure.

4.1.6 Results

We begin by reporting participant rates of perceived nuclear prominence in fillers, as an overall gauge of participants' attention during the prominence identification task. As stated above, the filler items (see example 4) each featured an instance of narrow contrastive focus which occurred in various positions within a sentence (initially, medially, and finally). Across participant groups, the mean rates of perceived nuclear prominence in relation to contrastive focus ranged between 0.87 and 0.91. In the Russian version of the task, Russian NSs chose the nuclear-accented word as prominent in 88% of the items, and English-Russian bilinguals – in 0.91%. In the English version of the task, English NSs chose the nuclear accented noun as prominent in 87% of the fillers. The obtained rates of perceived nuclear prominence, consistently high independent of the language of the task or the participant group, serve as evidence of on target, accurate performance by all participants.

Next, we examine the rate of perceived nuclear prominence in the test sentences each containing an instance of non-contrastive focus. We proceed by first reporting results obtained from the two groups of monolinguals (reference) speakers, and next - from the group of English-Russian bilinguals (see Table 3) for results summary.

The dependent measure in the data analyses for Experiment 1 is the likelihood of the nuclear pitch accented noun being rated prominent by the listeners. Inferential analyses modeling the likelihood of perceived nuclear prominence in the test sentences consisted of multinomial mixed-effects logistic regressions with *constituent order* (Russian data only), *cloze test score* (bilinguals' data only), and tonal measures of *pitch maxima*, *minima*, and *pitch excursion* entered as fixed effects. All tonal measures were coded separately for subjects and objects. The random effects for each model consisted of *participant* and *test item* (slopes and intercepts).

4.1.6.1 The English monolinguals

In the data obtained from the English monolinguals, the mean rate of perceived object prominence reached 0.08 (SD = 0.27) revealing an overwhelming preference to select the sentence-initial subject as prominent (mean = 0.73, SD = 0.45). Additionally, the verb was identified as prominent in approximately 20% of the test sentences (mean = 0.19, SD = 0.4). A mixed-effects multinomial logistic regression further revealed that, relative to the baseline category of the clause-final nuclear accented object, both the clause-initial subject and the clause-medial verb in the English stimuli were more likely to achieve perceived prominence, based on their acoustic-prosodic expression.

Among the tested acoustic-prosodic parameters, including the tonal measures, higher values of pitch minima over subjects and verbs ($z_{\text{subjects}} = -5.58$, $z_{\text{verbs}} = -4.63$, respectively, all p -values < 0.0001), as well as greater pitch excursion over these constituent categories ($z_{\text{subjects}} = -5.0$, $z_{\text{verbs}} = -3.66$, respectively, all p values < 0.0001), were negatively predictive of the nuclear accented object prominence, coded as base outcome in the regression model.

4.1.6.2 The Russian monolinguals

Here, we examine the rate at which object nominals in the SVO test sentences and sentence-final, nuclear-pitch accented subjects in the OVS test sentences, were identified as prominent by

monolingual Russian listeners. Because of a very low incidence of perceived verb prominence (<2% of all ratings) in the Russian monolinguals' data, the present discussion takes into account subject and object nominals only.

The mean rate of perceived nuclear prominence was 0.4 (SD = 0.5) in the SVO stimuli and 0.3 (SD = 0.46) in the OVS stimuli (overall task mean rate of nuclear prominence = 0.35, SD = 0.5). A mixed-effects logistic regression assessed the contribution of the acoustic-prosodic expression in subject and object nominals to their respective prominence rates. The fixed effects of interest included constituent order and the tonal measures of pitch. In this analysis, local pitch minima were excluded due to a collinearity effect. The pitch measures from subject and object nominals, which could be phrase-initial or phrase-final due to the constituent order manipulation, were coded separately and further interacted with constituent order.

The rate of perceived nuclear prominence was greater in the SVO sentences ($z = 2.14$, $p = 0.03$). In the SVO order, higher pitch maxima and excursion in sentence-final nominals were positively correlated with the probability of nuclear prominence (pitch maxima: $z = 1.99$, $p = 0.05$; pitch excursion: $z = 2.19$, $p = 0.03$).

4.1.6.3 The English-Russian bilinguals

This next set of results in the present experiment pertains to the performance of the English-Russian bilinguals. Because of a low incidence of perceived verb prominence (<3% of ratings) in the English-Russian bilinguals' data, the present discussion takes into account subject and object nominals only.

The mean rate of perceived nuclear prominence reached 0.45 (SD = 0.5). Differences in the mean rate of nuclear prominence due to constituent order appeared numerically low: SVO = 0.46 (SD = 0.5); OVS = 0.43 (SD = 0.5). A mixed-effects logistic regression evaluated the contribution of the acoustic-prosodic expression in subject and object nominals to their respective perceived prominence rate. Once again, local pitch minima were excluded due to collinearity. The model tested an additional main effect of participants' TL proficiency, as measured by means of a cloze test which all bilinguals speakers completed as a part of the present study.

The analysis returned no effect of constituent order. Both tested pitch measures (maxima and excursion size) yielded significant main effects (omitted for brevity) and interacted with

constituent order, as follows. In the SVO order, the size of pitch excursion (st) as well as pitch peak height over the sentence-final object were positively associated with the likelihood of the nuclear pitch prominence (excursion: $z = 4.72$; $p < 0.0001$; peak height: $z = 4.24$, $p < 0.0001$). The size of pitch excursion and peak height over the sentence-initial subject, on the contrary, were negatively associated with the likelihood of perceived nuclear pitch prominence (pitch excursion: $z = -4.14$, $p < 0.0001$; pitch maxima: $z = -2.15$, $p < 0.03$). Participants' performance on the multiple-choice cloze deletion score was positively, albeit weakly, predictive of how likely they were to select the nuclear accented noun as prominent, across the tested constituent orders ($z = 1.82$, $p = 0.07$).

4.1.6.4 Cumulative analysis of experiment 1 data

One final component of the present analysis is the model fit to the data obtained from all participants who completed the Russian version of the task, including the Russian monolinguals and the English-Russian bilinguals. The joint analysis revealed no main effect of language background or constituent order but highlighted the significant contribution of the tonal measures of nuclear prominence to perceived prominence ratings. The pitch peak height over sentence-initial subjects was negatively predictive of perceived nuclear prominence ($z = -4.42$, $p < 0.0001$), while higher pitch maxima over the sentence-final objects were positively associated with the likelihood of perceived nuclear prominence ($z = 5.04$, $p < 0.0001$). Similarly, all Russian-speaking participants were sensitive to the size of the pitch excursion over sentence-initial subjects ($z = -4.35$, $p < 0.0001$) and sentence-final objects ($z = 5.95$, $p < 0.0001$).

4.1.7 Discussion

Experiment 1 pursued the following questions: (1) What cues underly perception of prosodic prominence in simple transitive Russian sentences? (2) Do English-Russian bilinguals pattern with baseline Russian speakers in their prominence ratings?

In order to answer these questions, listeners provided ratings of perceived word-level prominence in the experimental stimuli based on acoustic-prosodic expression alone, i.e., in the absence of context cues. The experimental sentences were designed with the nuclear pitch accent on the sentence-final word, aligning with the preferred location of the main phrasal prominence in Russian and English.

TABLE 3 The mean rate of perceived nuclear prominence (means, SD) in the tested groups.

Prominent word category:	Sentence-final noun (nuclear accented)	Sentence-final noun by constituent order		Verb	Sentence-initial noun
English monolinguals	0.08 (0.27)			0.2 (0.4)	0.73 (0.45)
Russian monolinguals	0.35 (0.5)	SVO: 0.4 (0.5)	OVS: 0.3 (0.46)	<0.02	0.64 (0.4)
English-Russian bilinguals	0.45 (0.5)	SVO: 0.46 (0.5)	OVS: 0.43 (0.5)	<0.03	0.52 (0.5)

In a test sentence "The wolf scared the fox," sentence-final noun" refers to "fox," 'sentence-initial noun' refers to "wolf," and "verb" refers to "scared."
 "Sentence-final noun by constituent order" refers to objects in SVO and subjects – in OVS Russian stimuli sentences.

We predicted that all participants would demonstrate sensitivity to tonal measures contributing to pitch movement at the phrasal level in both English and Russian. Asymmetric rates of nuclear pitch prominence in SVO vs. OVS order were anticipated for the Russian monolinguals but not for the English-Russian bilinguals. This difference in prediction arises from the distinct grammars underlying each language: In Russian, phrasal prosody interfaces with constituent order in expressing information structure, leading constituent order to contribute to perceived word-level prominence [see Luchkina and Cole (2021) for a recent empirical investigation]. In contrast, in English, prosodic cues serve as the primary means of signaling prominence, while constituent order flexibility remains highly limited.

As predicted, all participants exhibited sensitivity to tonal measures in the modal speakers' read performance, including local pitch minima and maxima and the pitch excursion at word level. These cues supported near-ceiling rates of perceived nuclear prominence in filler sentences each featuring a contrastively accented word. These high rates of nuclear prominence in relation to the contrastive accenting patterns (LH* in Russian, L + H* in English) are consistent with recent research addressing contrastive focus in English (Bishop, 2012) and in Russian (Ionin et al., 2023b). The latter study, specifically, examined perceived contrastive focus prominence and identification in L1 and L2 Russian. Results reported by Ionin et al. (2023b) support that both Russian monolinguals and Russian-English bilinguals successfully identify contrastive foci in read recorded speech as prominent.

While this study presented fillers without supporting context, we attribute the high rates of perceived nuclear prominence in the filler sentences to the prosodic characteristics of the contrastively accented word. As reported in Bryzgunova (1980), contrastive focus in Russian receives a distinct prosodic contour, referred to as the non-neutral IK-2. Under the contour in question, the nuclear prominence may occur anywhere in the utterance, non-nuclear accents as well as pitch downstep tend to be eliminated, and the accented syllable is produced with particularly high intensity.

In a similar vein, Bishop (2012) argues for considerably greater prosodic prominence of contrastive focus (in comparison to non-contrastive focus) in English. An investigation by Cole et al. (2019) lends empirical support for this view. Cole & colleagues conducted a prominence rating task using recorded excerpts of connected English speech. The rate of perceived nuclear prominence (downstepped H*) in neutral intonation sentences in Cole et al.'s study reached approximately 0.3 and further reached approximately 0.5 in the sentences featuring an instance of narrow contrastive focus (L + H*).

The relatively higher prominence rates obtained by Cole et al. and in the present study may be attributed to the fact that in the former study, listeners were presented with stretches of connected discourse, whereas in the present study – utterance-long segments presented without context.

While acoustic-prosodic predictors continued to play a determinant role during prominence identification in the test items recorded with neutral intonation, most listeners were unlikely to select the nuclear accented noun as prosodically prominent. More specifically, the English monolinguals identified the phrase-initial subject nominal as prominent in 72% of the stimuli sentences and rated the verb as prominent in 20% of the test items.

The relatively high rate of perceived subject prominence in the English stimuli sentences could be attributed to several other

factors. Branigan et al. (2008) make a compelling argument for the special perceptual status of sentence subjects in English, due to their agentive role and animacy. Even though all subjects and objects in the experimental sentences were animate nouns, in the absence of context, some listeners possibly treated the stimuli sentences as instances of broad focus (i.e., all new information). As the information status remained constant across each test sentence, the grammatical function, in line with Branigan's proposal, could have further contributed to a prominent reading of the subject nominals.

At the same time, systematically reduced tonal measures in the vicinity of the phrase-final objects in the English stimuli have led to their relatively lower perceived prominence ratings. This proposal is further supported by the fact that (1) listeners were explicitly instructed to respond to the relative prosodic prominence at the word level during the prominence identification task and (2) phrase-finally, the tonal expression of pitch prominence is naturally acoustically reduced [see Katz and Selkirk (2011) and Yeung et al. (2019) for more discussion].

The relatively high rate of perceived subject prominence in the English stimuli sentences could be further attributed to several other factors. Branigan et al. (2008) make a compelling argument for the special perceptual status of sentence subjects in English, due to their agentive role and animacy. Even though all subjects and objects in the experimental sentences were animate nouns, in the absence of context, some listeners possibly treated the stimuli sentences as instances of broad focus (i.e., all new information). As the information status remained constant across each test sentence, the grammatical function, in line with Branigan's proposal, could have further contributed to a prominent reading of the subject nominals.

Results from the monolingual English speakers overlapped with those obtained from the Russian monolinguals in several ways. Specifically, both groups were responsive to the tonal measures in the model speakers' read performance and prioritized subject prominence over object prominence. Specifically, pitch excursion size predicted the likelihood of subject nominal prominence in both listener groups. Likewise, all monolinguals responded to the relative height of the pitch peaks over the phrase-final nominal constituents (as revealed in the joint analysis of the Russian task data) when selecting the prominent word.

The native-like perception of the tonal correlates of prominence in Russian may be attributed to positive transfer from the dominant language and, for some bilingual participants, to early exposure to Russian deemed critical for integrating phrasal prosody with the rest of the utterance, during listening (Laleko, 2024).

The impact of constituent order is where the performance of the two Russian speaking groups appeared to diverge. Specifically, the rate of perceived nuclear prominence in the Russian monolinguals' data was consistently higher, by 10% on average, in the SVO stimuli sentences compared to the OVS sentences. This difference stemmed from the tendency by the Russian monolinguals to select the sentence-initial object as prominent in the non-canonical OVS order. This result is consistent with the perception and production of alternate constituent orders in Russian previously investigated by Luchkina and Cole (2016, 2021). Luchkina and Cole (2016) reported evidence of comprehensive prosodic augmentation by Russian native speakers of words occurring *ex-situ*, such as the

sentence-initial object in the OVS sentences tested in the present study [see Vainio and Järvikivi (2006), Patil et al. (2008), and Luchkina et al. (2015) for similar findings in other flexible word order languages]. A follow up investigation by Luchkina and Cole (2021) found that *ex-situ* words in Russian are also more likely to be perceived as prominent during listening comprehension by adult native listeners.

In contrast to the results obtained from the monolingual speakers, the perception data from English-Russian bilinguals revealed a null effect of constituent order, despite satisfactory accuracy on the case check task and the cloze deletion test. Considering that both English and Russian utilize nuclear pitch prominence for marking focus, alternations in constituent order present a crucial asymmetry between the two languages and serve as a central axis for drawing a comparison between the speaker groups under investigation. We revisit the effect of constituent order on the rate of perceived nuclear prominence in the general discussion section 4.

4.2 Experiment 2: backward focus assignment

In Experiment 2, we tested if English-dominant Russian bilinguals can integrate nuclear pitch prominence and constituent order with discourse context to determine the word in sentence focus. The task is modeled after Breen et al. (2010) who previously investigated focus assignment in English, with functionally monolingual native English speakers.

4.2.1 Materials

Russian native speakers and proficient English-Russian bilinguals were compared on the use of the tonal focus correlates and linear ordering of sentence constituents during focus assignment.

Materials consisted of the 24 sentences tested in experiment 1 [see (8)], and 48 wh-questions [see (8.1) and (8.2)]. Each test sentence was paired with two wh-questions which cued the focal status of the subject or the object nominal in the test sentence. To illustrate, in the example (8) below, the object question in (8.1) correctly sets the nuclear pitch accented object ‘VOLKA’ (eng.: “wolf.GEN”) in focus; conversely, the subject question in (8.2) incorrectly sets the subject noun ‘LISA’ (eng.: “fox.NOM”) in focus.

(8)	Lisa Fox.Nom ‘The fox scared the wolf’	napugala scared	VOLKA wolf.Acc	[SVO order]
(8.1)	Kogo whom ‘Whom did the fox scare?’	napugala scared	lisa? fox.Nom	[object question]
(8.2)	Kto who ‘Who scared the wolf?’	napugal scared	volka? wolf.Acc	[subject question]

The filler sentences from experiment 1 were paired with two wh- or yes-no questions, in the same format as the test items.

4.2.2 Participants

The Russian-speaking participants who completed the auditory prominence identification task continued to experiment 2, including the Russian monolinguals (the reference group) and the English-Russian bilinguals (the test group).

4.2.3 Procedure

In each trial, listeners had to decide whether each test sentence was a response to an object or a subject question to determine which of the two nouns, the subject or the object, the speaker intended as the sentence focus. They were instructed to select a context question which best matched the target sentence, using the two provided options. One of the options set the nuclear accented word in the target sentence in sentence focus (Match), while the other– assumed a focal reading for a non-nuclear accented noun (Mismatch).

The experimental sentence was presented auditorily and the 2 context questions were presented side by side, in writing. Participants listened to the target sentence and selected the matching question with a mouse click.

4.2.4 Testable predictions

In the backward focus assignment task, we investigate the rate at which new information foci, nuclear pitch accented in the sentence-final position, were successfully disambiguated by listeners, as indicated by the rate of choosing matching contexts over non-matching ones. The most salient cues to sentence focus made available in the sentence stimuli included constituent order and tonal correlates in phrase-final subject and object nominals.

Because information structure serves as an interpretative license for constituent re-ordering in Russian, we predicted that the Russian monolinguals would exhibit preference to assign focus to the nuclear-accented noun sentence-finally, across the tested constituent orders. This same prediction can be further extended to the bilinguals’ group if bilinguals at higher TL proficiency successfully associate word order with distinctions in the information structure. If, on the contrary, an effect of constituent order emerges in the bilinguals’ data, it would be indicative of transfer from the dominant grammar, where the said effects of information structure on constituent order are not found.

Despite the lower rates of perceived phrase-final prominence obtained during nuclear prominence identification, we nevertheless anticipated above-chance rates of matching context-answer pairings due to the qualitatively different nature of the task at hand. While not instructed to attend closely to the intonation in the target sentences, participants were expected to perceive the stimuli more holistically and take information structure in the question-answer pairs into an account. Critically, this expectation holds both for English-dominant and Russian monolingual speakers, owing to an overlap in (1) the default, phrase-final placement of the nuclear pitch prominence in both these languages and (2) the tonal correlates of the nuclear pitch prominence marking non-contrastive focus in Russian and English.

4.2.5 Results

As with Experiment 1, we first report the rate of focus assignment in the filler items (see example 4) which featured an instance of narrow contrastive focus occurring in various positions within a sentence (initially, medially, and finally). Native Russian speakers’ target focus assignment rate reached an average of 0.77 and varied among the different types of fillers (range: 0.70–0.83). English-Russian bilinguals

demonstrated comparable performance, at the average rate of on target contrastive focus assignment of 0.82. These results support that all participants paid attention and understood the task instructions.

The overall rate of on target focus assignment in non-contrastive focus items reached 0.63 (SD = 0.48). The Russian monolinguals achieved the mean accuracy of 69.4 (SD = 0.46). As shown in Figure 3, the mean accuracy rate was numerically higher on SV[O]_F items (mean = 0.74, SD = 0.44) than on OV[S]_F items (mean = 0.65, SD = 0.48). The English-dominant bilinguals were 0.56 accurate overall (SD = 0.5). Their accuracy also differed across the tested constituent orders: SVO: 0.59 (SD = 0.49); OVS: 0.53 (SD = 0.5).

The constituent order fixed effect in the monolingual speakers' data did not reach significance when evaluated in a mixed-effects logistic regression. The same analysis additionally assessed the contribution of the tonal focus correlates to the likelihood of accurate backward focus assignment, as well as their interaction with constituent order. The tonal measures of pitch maxima and excursion size in object nominals were positively predictive of accurate focus assignment to the nuclear pitch accented word under the SVO constituent order (object pitch excursion: $z = 2.27$, $p = 0.02$; object pitch maxima: $z = 1.87$, $p = 0.06$) by the Russian monolingual speakers.

An analogous model fit to the English-Russian bilinguals' data revealed significant main effects of constituent order ($z = 1.97$, $p = 0.05$), as displayed in Figure 4 and cloze test score ($z = 2.5$, $p = 0.01$). Furthermore, pitch peak height over the subject nominals

negatively predicted the accurate choice of the context question ($z = -1.99$, $p = 0.05$) in the SVO order.

The final component of Experiment 2 inferential analyses is a mixed effects model fit to the entirety of the Russian speakers' data, with participants' dominant language entered as a fixed effect. The model returned a significant main effect of participants' dominant language ($z = 2.22$, $p = 0.03$), reflective of a more accurate performance of the Russian monolinguals on backward focus assignment. The tonal measures extracted from the object nominal interacted with constituent order (pitch excursion: $z = 1.95$, $p = 0.05$; pitch maxima: $z = 2.16$, $p < 0.03$;) and predicted accurate focus assignment to sentence-final objects in SV[O]_F test sentences. The size of pitch excursion over the sentence-initial subjects was negatively predictive of the likelihood of focus assignment to the nuclear-accented noun in SV[O]_F test sentences ($z = -2.35$, $p = 0.02$). Finally, the effect of constituent order approached significance ($z = 1.8$, $p = 0.07$), driven by the more robust effect obtained in the bilingual participants' data but lacking in the data from the baseline speakers.

4.2.6 Experiment 2: discussion

In Experiment 2, we examined whether English-Russian bilinguals demonstrate alignment with native Russian speakers in their use of prosodic features and constituent order when assigning focus at the phrasal level. To this end, participants completed a backward focus

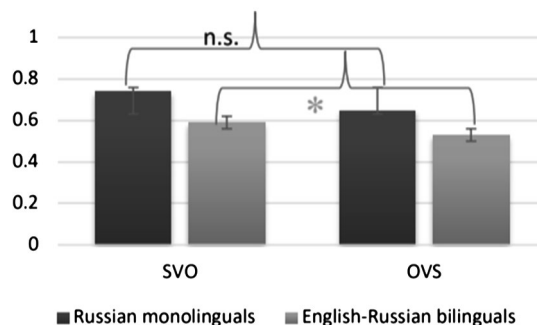


FIGURE 3

Mean accuracy rate (y-axis) obtained in the backward focus assignment task in the Russian stimuli sentences. Error bars represent Standard Deviation.

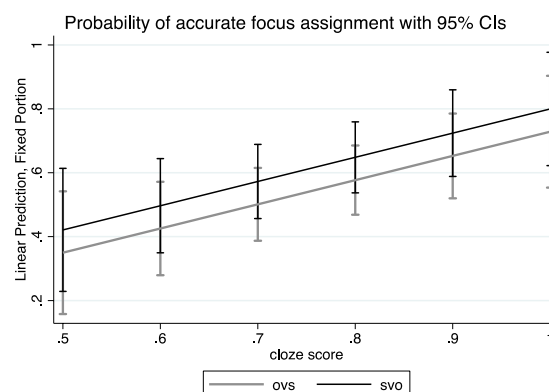


FIGURE 4

Likelihood (linear prediction) of accurate focus assignment in bilingual speakers' data. Top line: SVO stimuli sentences; lower line: OVS stimuli sentences. Error bars represent Standard Deviation.

assignment task in which they were presented with target sentences alongside two context options. The position of the nuclear pitch-accented word in each target sentence was invariably sentence-final. Listeners were tasked with selecting the context question that would accurately place the sentence-final nominal in focus.

We predicted the Russian monolinguals to rely on the acoustic-prosodic cues, in combination with constituent order, when performing backward focus assignment. As long as both tested constituent orders conform to the same interpretative license, the listeners should anticipate sentence focus phrase-finally and in alignment with the default, phrase-final nuclear prominence lending site. This logic supports the expectation of comparable focus assignment accuracy across the tested constituent orders. This prediction, if borne out, goes against the observed effect of constituent order in the monolinguals' prominence identification results obtained in Experiment 1. Furthermore, we expected English-Russian bilinguals to demonstrate performance above chance levels. This prediction is supported by satisfactory assessment outcomes of their proficiency in the TL, as well as the transfer of pitch marking for sentence focus from the dominant language.

Listeners demonstrated sensitivity to the tonal correlates of sentence focus, which systematically contributed to the choice of the matching context for both groups. Different listening patterns emerged, whereby the Russian monolinguals attended to the tonal properties of object nominals, such that greater pitch excursion cued object focus, irrespective of constituent order, and higher pitch peaks further supported object focus assignment under the baseline SVO order. In contrast, English-Russian bilinguals closely tracked the relative pitch prominence of the sentence-initial subject nominals in the SVO order, such that their pitch peak height was inversely associated with the likelihood of sentence-final object focus.

We interpret the listening pattern of English-Russian bilinguals to be influenced by transfer from their dominant language. In English, where constituent order flexibility is limited, it is more likely for the nuclear prominence to occur in variable phrasal positions, i.e., non-utterance-finally. This may have led our bilingual listeners to anticipate, on a probabilistic basis, a non-sentence-final nuclear pitch accent in the stimuli sentences. While this interpretation is tentative, it aligns with the higher rate of assigned subject foci in SVO order in the English-Russian bilinguals' data (0.4), compared to the lower rate (0.26) – in the Russian monolinguals' data. It's noteworthy that Russian monolinguals, too, remained open to the possibility of subject focus in the SVO order but used pitch excursion (rather than peak height) as a leading tonal correlate supporting subject focus assignment to the sentence-initial subject. This unified analysis of the backward focus assignment data from both participant groups further supports the contribution of the tonal correlates of nuclear prominence in Russian, positively predicting accurate focus assignment across both tested constituent orders.

The performance of the two groups diverged when considering the effect of constituent order systematically varied in the stimuli sentences. All participants appeared more inclined to assign focus to the clause-final object nominal under the baseline SVO order. While the difference in the mean rates of accurate focus assignment in SVO vs. OVS stimuli sentences was numerically greater in the monolingual participants' data, it did not reach significance.

The null effect of constituent order in the monolinguals' data aligns with our prediction, indicating that the preferred strategy for monolingual listeners was to assign focus to the phrase-final nominal, which naturally aligns with the nuclear prominence lending site in

Russian (Neeleman and Titov, 2009). This interpretation is critically supported by the fact that the Russian monolinguals exhibited a preference to assign focus to the sentence-final subject in the OVS stimuli sentences.

On the contrary, English-Russian bilinguals consistently showed a preference for assigning focus to the sentence-final object nominal in the SVO stimulus sentences. Simultaneously, they demonstrated a higher rate (47%) of assigning focus to the object in the OVS order. As expected, the bilinguals' distinct approach to focus assignment was mirrored in their performance on the cloze deletion test, utilized as a gauge of target language proficiency. Specifically, bilinguals who performed well on the cloze test were more likely to assign focus to the nuclear accented word.

Adding further support to the qualitatively different approach to focus assignment in the two groups, a significant main effect of dominant language emerged in the unified analysis of the data, indicating an overall stronger tendency among the Russian monolinguals to assign focus to the sentence-final nominal constituent, irrespective of the constituent order. These findings support the prediction that Russian monolingual speakers relied on constituent order as a heuristic during focus assignment more than the English-dominant bilinguals.

5 General discussion

The present study investigates bilingual competence in the domain of information structure in Russian, a free word order, pitch-accenting language. The primary focus is on the simultaneous use of intonational prominence and constituent order as means of encoding sentence focus by English-Russian bilinguals with English as their primary or dominant language. The population of interest has previously demonstrated varied acceptability of non-SVO orders in Russian (Laleko, 2022; Ionin et al., 2023a) in conjunction with non-target-like perception of prosodic prominence used to mark non-contrastive narrow focus in object-final and subject-final transitive sentences (Laleko, 2024; Luchkina et al., in press).

The added complexity in relation to non-contrastive new information focus in Russian motivates the analysis of reference data from adult monolinguals commanding each of the languages of our bilingual participants. In the present study, this leads us to include monolingual Russian and English speakers whose auditory perception data and focus assignment data are used to establish baseline against which we then compare the results from the bilinguals.

The first listening task tested participants' perception of the main phrasal prominence in a series of simple transitive sentences. Given the often-subtle nature of the acoustic-prosodic cues in the expression of a phrase-final nuclear pitch accent, we aimed to determine if listeners perceive the accented word as prominent based on its tonal expression. Additionally, we investigated whether the perception of phrasal prominence is influenced by the linear order of sentence constituents in Russian, in comparison to a fixed constituent order in English. Because a significant contribution of discourse context toward prominence identification has been previously reported by Luchkina et al. (in press), we chose to center the present investigation on the auditory perception of prominence in the absence of context cues.

The rate of perceived nuclear prominence was significantly lower (<10% of all ratings) in the data from monolingual English speakers

who readily rated the sentence-initial nominal or the verb as more prosodically prominent than the sentence-final pitch-accented nominal. As far as the prosodic expression is concerned, the low rate of sentence-final nuclear prominence in the English stimuli sentences can be attributed to the frequent occurrence of vocal fry in the vicinity of the sentence-final object which often compromised the realization of the intended pitch contour and rendered the sentence-final object less prosodically prominent compared to the rest of the sentence.

In Russian, analyses of the tonal measures in the sentence stimuli revealed that the phrase-final noun, despite being nuclear pitch-accented, appeared less prosodically expressive compared to the non-phrase-final material. Nevertheless, in comparison to the English reference group, the Russian monolinguals were five times more likely to perceive the sentence-final, nuclear accented nominal as prominent in the baseline SVO stimuli sentences and nearly four times more likely—in the subject-final OVS order. Results from the English-Russian bilinguals did not align fully with either monolingual reference group. First, there was a very low incidence of verb prominence, unlike in the English monolinguals' data. Second, bilinguals' ratings were unaffected by constituent order in the test sentences, unlike in the Russian monolinguals' data. Numerically, bilingual listeners were more likely to rate the nuclear-pitch accented noun as prominent, albeit the overall rate of perceived nuclear prominence remained under 50%.

The relatively lower rates of perceived nuclear prominence obtained in all participant groups appear even more notable considering that both Russian and English are known to default to phrase-final nuclear pitch prominence (Bryzgunova, 1980; Beckman and Pierrehumbert, 1986) and warrant further analysis of the individual contribution of tonal cues to nuclear pitch prominence in each listener group. At the same time, these results point to the probabilistic nature of perceived prosodic prominence during listening comprehension. The term “probabilistic” in the context of the prominence identification task administered in the present study translates into notable levels of individual variation in perception, stemming from distinct approaches to prominence identification adopted by linguistically naïve listeners. This variability suggests that some of the listeners tested in the present study prioritized prosodic expression as the primary “pathway” to prominence, while others relied on grammatical function, information status, and other discourse cues [see Branigan et al. (2008), Watson (2010), and Cole (2015) for further discussion]. Although all listener groups were explicitly instructed to focus on prosodic cues to determine prominence, some may have, in principle, evaluated perceived information prominence holistically, i.e., considered fundamentally non-prosodic cues to prominence, as discussed here and in section 3 above.

Our results align with a recent English study by Yeung et al. (2019) who established a largely probabilistic mapping between the cues used by L1 English speakers to express focus in elicited production and by listeners - during auditory comprehension of recorded speech. Similar findings on Russian were reported by Luchkina and Cole (2016) in an investigation of prosodic prominence correlates in read recorded speech by 15 native Russian speakers. The study found that several speakers failed to prosodically augment words which were referentially new in read discourse — a finding which parallels that of Yeung et al.'s (2019) study on English.

The probabilistic nature of perceived nuclear prominence in the present study may be further attributed to the fact that our stimuli sentences were presented without context against which the nuclear status of the pitch accent over the sentence-final noun could be interpretatively validated. As a result, listeners may have developed different heuristics leading to distinct prominence percepts. For example, some respondents may have been sensitive to phrasal prosody, which includes downstepping in pitch across an utterance. This feature supports greater perceptual prominence for words occurring earlier in the string, as opposed to the nuclear-accented, sentence-final word. This interpretation is supported by the high rate of perceived prominence associated with the sentence-initial noun reported by most listeners. Other participants may have relied on known information structural templates shared by Russian and English, since in both these languages, discourse-given information tends to be placed early in the utterance, while discourse-new information often appears at the end. This pattern supports a prominent reading of the utterance-final, nuclear-accented noun.

By considering these different heuristics, we can better understand the variability in listeners' prominence percepts. Findings of Experiment 1 lead us to propose that a unity of prosodic cues and discourse heuristics is what may be necessary for a full-fledged percept of nuclear prominence to emerge within a listener. To test this proposition, one may require conducting an additional task asking the listeners to point out the most prominent word, as opposed to the most prosodically prominent word, while making the context available. A similar design has been previously implemented in Luchkina et al. (in press) where it gave rise to prosodic correlates and discourse cues to prominence being co-interpretable by listeners.

In Luchkina et al. (in press), Russian monolinguals and English-Russian bilinguals were tested using brief question-answer exchanges. The question sentences systematically set either the subject or the object in the SVO and OVS targets in sentence focus. The same target sentences were used as in the present study, giving rise to pragmatically felicitous and non-felicitous exchanges. In the non-felicitous items, the question sentence placed focus on the sentence-initial noun, which clashed with the sentence-final nuclear prominence in the answer sentence.

By manipulating context felicity, Luchkina et al. found that in felicitous question-answer pairings, the rate of nuclear prominence was greater, due to context unambiguously reinforcing the pitch accenting status of the sentence-final noun. The non-felicitous context, however, set a non-nuclear accented word in focus and thereby made listeners less likely to rate the nuclear-accented word as prominent. Despite the more variable proficiency levels of the English-Russian bilinguals tested by Luchkina et al. (in press), a robust effect of context felicity emerged, highlighting the tight interplay between context and prosodic cues in the perception of phrasal prominence. For instance, the rate of nuclear prominence in OVS targets dropped by 40% when a non-felicitous context was provided.

Since the present study investigates the role of intonational cues to prominence, discourse context was made unavailable. As a result, participants demonstrated greater reliance on prosodic expression.

Despite the overall lower incidence of perceived nuclear prominence reported in Experiment 1, all listener groups were responsive to word-level acoustic-prosodic tonal expression, which often rendered a non-phrase-final element prosodically distinct, even in the absence of a nuclear accent. More specifically, all listeners demonstrated sensitivity to local pitch maxima and a relative size of pitch excursion, independent of the language of the task. Numerically

comparable rates of perceived nuclear pitch prominence obtained from the bilingual raters provide evidence of successful transfer in the domain of phrasal prosody from the dominant language (English), even though the tonal signatures of nuclear accents in the English (H*) and Russian (HL*) stimuli sentences were prosodically distinct.

We conclude that, phrase-finally, the nuclear pitch-accented status might not be as straightforward from a perceptual perspective, particularly in the absence of context cues that delineate the information structure of the utterance at hand and may further enhance the prominent status of the sentence-final word.

Bishop (2012) is an illustrative empirical study which further supports this proposal using data from English. In the prominence rating task administered by Bishop (2012), listeners were presented with spliced productions of SVO sentences in which the entire VP was in broad focus (both verb and object were accented) paired with contexts which only supported the focal status of the object. Following this manipulation, the listeners reported hearing the object as more prominent than the verb, and thereby revealed a robust effect of context cues and listener-based expectations of prosodic prominence in connected discourse.

We conclude that the construct of word-level prominence prominence extends beyond the specific focus of this study and is both conceptually and perceptually broader than the dependent measure examined in Experiment 1 (i.e., the rate of perceived nuclear prominence) and its critical predictors (tonal measures of nuclear accent and phrasal placement of the prominent word). Further investigations into the psychological, expectation-driven nature of perceived prominence coincidental with new information focus in various phrasal positions in Russian are warranted, given the probabilistic relationship between production-perception data reported in the present study and related work.

Experiment 2 asked listeners to identify the word in each test sentence that could plausibly serve as the focus exponent and provided listeners (albeit in a reverse fashion) with two contexts to choose from. By offering context alternatives, the the focus assignment task altered the participants' listening experience and made them more likely to perceive nuclear prominence through the lens of the provided discourse scenarios. It is noteworthy that both tested groups responded to the task similarly, by accepting the focal reading of the nuclear-accented noun for the majority of the stimulus sentences (cf. Russian monolinguals: 70%; English-Russian bilinguals: 56%).

Of particular interest to the present study is how prosodic prominence identification and focus assignment proceed in the subject-final OVS order in Russian. Previously, Hoot (2017) reported lower acceptability of phrase-final representational subject foci (OV[S]_F), in comparison to phrase-initial ([S]_FVO), by native and heritage speakers of Mexican Spanish. Kallestinova (2007), however, reported an opposite pattern of relative acceptability in native Russian speakers who preferred the subject-final order when the subject was discourse-new. This preference may be accounted for by the fact that placing the subject nominal phrase-finally aligns it with the default location of the main phrasal prominence and, by doing so, effectively signals its focal status during listening.

An opposite pattern of acceptability judgments in Russian was recently discussed in Laleko (2024), where 14 baseline speakers equally accepted the prosodic encoding of sentence foci occurring *in situ*, as well as focus encoding via constituent reversal, resulting in the alignment of the focused word with the nuclear prominence phrase-finally (p.15). This duality of

strategies available for focus marking led us to further investigate the integration of constituent order with prosodic cues during prominence identification and focus assignment.

In the present study, notable differences emerged when comparing the effects of constituent order in bilingual and monolingual groups during both prominence identification and focus assignment. While monolingual participants were more likely to identify the nuclear accented noun as prominent under the baseline SVO order, bilinguals did not display such a preference, despite the overlap in the baseline order in the dominant grammar and the TL grammar.

Notwithstanding the lower rates of perceived nuclear prominence obtained in the first listening task, we anticipated more accurate performance in the focus assignment task to be achieved via integration of the constituent order with the prosodic expression in the test sentences. Critically, this prediction was confirmed in the data from monolingual Russian speakers, who were more likely to treat the sentence-final nominal constituent as focal in either constituent order, viewing constituent order as a means of focus expression. In contrast, our bilingual participants consistently preferred assigning focus to the nuclear accented noun in the SVO order alone.

These results support recent research on Russian L2 learners and heritage speakers (e.g., Ionin and Luchkina, 2019), which documented a common tendency toward non-target-like interpretations of non-canonical OVS sentences by both adult L2 learners and heritage speakers of Russian. In a similar vein, Ionin et al. (2023a) and Laleko (2022) observed lower overall acceptability of OV[S]_F sentences compared to the baseline SV[O]_F order, even in the presence of felicitous discourse context.

To date, the sole published investigation addressing the acquisition of Russian focus using methods comparable to ours is Ionin et al. (2023b). Ionin and colleagues tested listeners' identification of nuclear prominence in relation to contrastive focus in Russian. This study reported highly successful, native-like performance by English-dominant Russian bilinguals, likely attributed to the distinct nature of the nuclear pitch accent marking contrastive focus in Russian. Critically, in Ionin et al. (2023b), both early and late English-Russian bilinguals demonstrated native-like acceptability and perception of Russian contrastive focus prosody tested under variable constituent orders.

In contrast to Ionin et al. (2023b) findings, our bilingual participants demonstrated dissimilar focus assignment preferences when compared to the baseline monolingual Russian speakers. Specifically, when the order of nominal constituents in the test sentences was reversed, bilinguals were notably less inclined to treat the nuclear-accented noun as focus. This suggests that under a non-canonical constituent order, bilinguals differ in their focus assignment preferences from the baseline speakers who clearly identify the nuclear-accented noun as focus, across the tested constituent orders. Further evidence supporting the non-native-like performance of the bilingual speakers was observed through the influence of the TL proficiency observed in both experimental tasks. Specifically, better performance on the cloze deletion test predicted a higher likelihood of selecting the nuclear-accented noun as prominent in Experiment 1, as well as the focus exponent in Experiment 2.

The divergent perception patterns observed in our bilingual speakers' lend support to the Interface Hypothesis (Sorace and Filiaci, 2006), which predicts non-target-like acquisition patterns specifically at external interfaces. Sorace and Serratrice (2009) further discuss several factors that

contribute to the vulnerability of interface phenomena. These include underspecification of interface conditions in the native grammar, cross-linguistic influence, target language (TL) input (quality and quantity), and processing limitations. While the present study does not measure TL input or the processing resources required for focus assignment, the vulnerability in acquiring the Russian focus structure by adult English-Russian bilinguals may be attributed to the unique role of constituent order in signaling information structure and its linkage with phrasal prosody in Russian. Since neither of these properties are present in the L1 grammar, they remain underspecified and thus a potential source of cross-linguistic influence. Additionally, the three-way nature of the interface in question introduces greater inherent complexity, as it involves integrating prosodic cues indicative of the new information status of a discourse referent with constituent order during discourse processing, potentially exacerbating non-target-like performance.

Our approach to participant inclusion on the basis of a cut off TL proficiency score does not enable us to assess if the non-target-like performance observed in the present sample would generalize to speakers with even higher Russian proficiency and/or exceptionally early naturalistic exposure to the TL, similar to studies conducted by [Laleko \(2022\)](#) and [Ionin et al. \(2023c\)](#). The study by [Ionin et al. \(2023c\)](#) found that bilinguals with early naturalistic exposure to Russian in a heritage setting were more native-like in accepting the OV[S]_F order in response to narrow subject focus in comparison to late L2ers. Additionally, early bilinguals tested in [Ionin et al.](#) were more likely to accept the OVS order as their TL proficiency increased. This trend was not observed among late learners, regardless of their proficiency levels. [Laleko \(2022\)](#), similarly, discovered that English-Russian bilinguals representing diverse backgrounds and proficiency levels generally exhibited lower acceptability of the OVS order, except for high-proficiency heritage speakers who displayed greater acceptability of transitive OV[S]_F sentences. While bilinguals in the present study were chosen based on their TL proficiency rather than their age of exposure to Russian (due to an imbalanced number of speakers with early vs. late AOE), it is conceivable that a more on target performance could emerge in a homogeneous sample of high-performing listeners with particularly early ages of exposure to the TL.

6 Conclusion

This study offers new, data-driven insights into the acquisition of the relationship between prosody, constituent order, and information structure in Russian. Our experimental methodology systematically compares how new information focus is signaled in participants' dominant language (English) and the TL (Russian).

We reported that, in the absence of discourse context, the nuclear pitch accent aligned with the phrase-final nominal acts as a probabilistic, rather than deterministic, indicator for a prosodically prominent reading of the accented word. Further supporting this finding, the evidence of acoustic-prosodic augmentation in relation to the nuclear pitch accenting was subtle in the production data of our model speakers, across tested languages.

As we investigated the link between sentence-final nuclear prominence in Russian and the focal interpretation of clause-final nominal constituents, we discovered notable differences between Russian monolingual speakers and English-Russian bilinguals. Unlike bilinguals,

Russian monolinguals exhibited a stronger expectation for phrase-final nuclear prominence in the SVO constituent order compared to the reversed OVS order. Conversely, during focus assignment, English-dominant bilinguals were inclined to assign new information focus to the nuclear accented nominal in the SVO order, and less so - in the subject-final OVS order. In contrast, Russian monolinguals' preference to assign focus to the nuclear accented nominal upheld irrespective of the constituent order.

Varying performance patterns among baseline Russian speakers and English-Russian bilinguals reveal two key findings: (1) there are no clear a priori expectations about where the main phrasal prominence will occur within an utterance in either English or Russian, and (2) there is less certainty in integrating non-canonical constituent order with phrasal prosody during focus assignment in Russian. This uncertainty leads bilingual listeners to consider both ex-situ elements—the sentence-initial object and the sentence-final subject—as plausible focus exponents.

These findings are largely in line with the IH, which predicts acquisition difficulties for language external interfaces ([Sorace and Filiaci, 2006](#); [Montrul and Polinsky, 2011](#)). However, they also highlight the need for testing additional samples of bilingual speakers with earlier ages of target language exposure in a naturalistic setting, as prompted by recent work by [Laleko \(2022\)](#) and [Ionin et al. \(2023c\)](#). Additional research with monolingual Russian speakers is also warranted, to further explore the division of labor between constituent order, prosodic expression, and information structural distinctions in focus marking in Russian.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Central Connecticut State University Institutional Review Board Committee. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

TL: Conceptualization, Writing – original draft, Writing – review & editing. TI: Conceptualization, Data curation, Investigation, Methodology, Project administration, Writing – review & editing. MG: Conceptualization, Data curation, Investigation, Methodology, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2024.1363980/full#supplementary-material>

References

- Bailyn, J. F. (1995). A configurational approach to Russian "free" word order: Cornell University.
- Beckman, M. E. (1996). The parsing of prosody. *Lang. Cogn. Proc.* 11, 17–68. doi: 10.1080/016909696387213
- Beckman, M. E., Hirschberg, J., and Shattuck-Hufnagel, S. (2005). "The original ToBI system and the evolution of the ToBI framework" in Prosodic typology: The phonology of intonation and phrasing. ed. S.-A. Jun (Oxford: Oxford University Press), 9–54.
- Beckman, M. E., and Pierrehumbert, J. B. (1986). Intonational structure in Japanese and English. *Phonology* 3, 255–309. doi: 10.1017/S095267570000066X
- Bishop, J. (2012). "(2012). Information structural expectations in the perception of prosodic prominence" in Prosody and meaning (trends in linguistics). eds. G. Elordieta and P. Prieto (Berlin: Mouton de Gruyter), 239–270.
- Bishop, J., Kuo, G., and Kim, B. (2020). Phonology, phonetics, and signal-extrinsic factors in the perception of prosodic prominence: evidence from rapid prosody transcription. *J. Phon.* 82:100977. doi: 10.1016/j.jwocn.2020.100977
- Bivon, R. (1971). Element order. Studies in the Modern Russian Language 7. London: Cambridge University Press.
- Boersma, P., and Weenink, D. (2024). Praat: doing phonetics by computer [computer program]. Version 6.4.17. Available at: <http://www.praat.org/> (Accessed August 8, 2024).
- Branigan, H. P., Pickering, M. J., and Tanaka, M. (2008). Contributions of animacy to grammatical function assignment and word order during production. *Lingua* 118, 172–189. doi: 10.1016/j.lingua.2007.02.003
- Breen, M., Fedorenko, E., Wagner, M., and Gibson, E. (2010). Acoustic correlates of information structure. *Lang. Cogn. Proc.* 25, 1044–1098. doi: 10.1080/01690965.2010.504378
- Bryzgunova, E. A. (1980). "Intonacija." In Russkaja grammatika, N. Ju Svedova, 96–122. Moskva: Nauka.
- Büring, D. (2009). "Towards a typology of focus realization. Information structure" in Information structure: Theoretical, typological and experimental perspectives. eds. M. Zimmermann and C. Fery (Oxford, UK: Oxford University Press), 177–205.
- Calhoun, S. (2010). The centrality of metrical structure in signaling information structure: a probabilistic perspective. *Language* 86, 1–42. doi: 10.1353/lan.0.0197
- Chafe, W. L. (1976). "Givenness, contrastiveness, definiteness, subjects, topics, and point of view" in Subject and topic. ed. C. Li (New York: Academic Press), 25–55.
- Chomsky, N., and Halle, M. (1968). The sound pattern of English. New York: Harper Row.
- Cole, J. (2015). Prosody in context: a review. *Lang. Cogn. Neurosci.* 30, 1–31. doi: 10.1080/23273798.2014.963130
- Cole, J., Hualde, J. I., Smith, C. L., Eager, C., Mahrt, T., and de Souza, R. N. (2019). Sound, structure and meaning: the bases of prominence ratings in English, French and Spanish. *J. Phon.* 75, 113–147. doi: 10.1016/j.jwocn.2019.05.002
- Cole, J., and Shattuck-Hufnagel, S. (2016). New methods for prosodic transcription: capturing variability as a source of information. *Lab. Phonol.* 7:8. doi: 10.5334/labphon.29
- Cruschina, S. (2022). Focus and focus structures in the romance languages. *Oxford Res. Encycl. Rom. Ling.*, 1–48. doi: 10.1093/acrefore/9780199384655.013.649
- Fletcher, J., and Evans, N. (2002). An acoustic phonetic analysis of intonational prominence in two Australian languages. *J. Int. Phon. Assoc.* 32, 123–140. doi: 10.1017/S0025100302001019
- Fruit, M. (2006). "L2 acquisition of focus constructions in European Portuguese and the L2 status of the syntax discourse interface" in Proceedings to the 8th generative approaches to second language acquisition conference (GASLA 2006) (The Banff Conference), 41–50.
- Gussenhoven, C. (2004). The phonology of tone and intonation. Cambridge: Cambridge University Press.
- Gussenhoven, C., and Rietveld, A. C. M. (1988). Fundamental frequency declination in Dutch: testing three hypotheses. *J. Phon.* 16, 355–369. doi: 10.1016/S0095-4470(19)30509-1
- Hertel, T. J. (2003). Lexical and discourse factors in the second language acquisition of Spanish word order. *Second. Lang. Res.* 19, 273–304. doi: 10.1191/0267658303sr2240a
- Hoot, B. (2017). Narrow presentational focus in heritage Spanish and the syntax-discourse interface. *Ling. Approach. Bilingual.* 7, 63–95. doi: 10.1075/lab.14021.hoo
- Ionin, T., Goldshtein, M., and Luchkina, T. (2023c). Word order in heritage Russian under different dominant languages. *J. Slavic Ling.* 31, 1–15.
- Ionin, T., Goldshtein, M., Luchkina, T., and Styrina, S. (2023a). Who did what to whom, and what did we already know? Word order and information structure in heritage and L2 Russian. *Ling. Appr. Bilingual.* 13, 343–371. doi: 10.1075/lab.20045.ion
- Ionin, T., and Luchkina, T. (2019). Scope, syntax and prosody in Russian as a second or heritage language, Exploring interfaces: Lexicon, Syntax, Semantics and Sound, 141–70. Cambridge: Cambridge University Press, Cambridge.
- Ionin, T., Luchkina, T., and Goldshtein, M. (2023b). Contrastive focus is acquirable: an investigation of Russian contrastive focus with English/Russian bilinguals. *Second. Lang. Res.* 40, 709–737. doi: 10.1177/02676583231178101
- Jasinskaja, K. (2016). Information structure in Slavic. *Oxford Handbook Inform. Struct.*, 709–732. doi: 10.1093/oxfordhb/9780199642670.013.25
- Kalvestinova, E. D. (2007). Aspects of word order in Russian. Ph.D. Thesis.: The University of Iowa.
- Katz, J., and Selkirk, E. (2011). Contrastive focus vs. discourse-new: evidence from phonetic prominence in English. *Language*, 771–816.
- Ladd, D. R. (2008). Intonational phonology. Cambridge: Cambridge University Press.
- Laleko, O. (2022). Word order and information structure in heritage and L2 Russian: focus and unaccusativity effects in subject inversion. *Int. J. Biling.* 26, 749–766. doi: 10.1177/13670069211063674
- Laleko, O. (2024). The complexity of word order change in a flexible system: on stability and variation in heritage Russian word order. *Formal Approach. Complex. Heritage Lang. Gramm.* 73.
- Luchkina, T., and Cole, J. S. (2016). Structural and referent-based effects on prosodic expression in Russian. *Phonetica* 73, 279–313.
- Luchkina, T., and Cole, J. S. (2021). Perception of word-level prominence in free word order language discourse. *Lang. Speech* 64, 381–412. doi: 10.1177/0023830919884089
- Luchkina, T., Cole, J. S., Jyothi, P., and Puri, V. (2015). "Prosodic and structural correlates of perceived prominence in Russian and Hindi" in Proceedings of the 2015 meeting of ICPHS.
- Luchkina, T., Ionin, T., and Goldshtein, M. (in press). "Word order and prosody effects in perception of sentence focus in Russian by native speakers and adult L2 learners" in Proceedings of FASL 2021.
- Luchkina, T., Ionin, T., Goldshtein, M., and Lysenko, N. (manuscript in preparation). Correlates of perceived focal prominence in Russian.
- Luchkina, T., Ionin, T., Lysenko, N., Stoops, A., and Suvorkina, N. (2021). Evaluating the Russian language proficiency of bilingual and second language learners of Russian. *Languages* 6:83. doi: 10.3390/languages6020083
- Makarova, E. N. (2021). Translating communicative intention and meaning from English into Russian. *Filologicheskij klass* 26, 132–143. doi: 10.51762/1FK-2021-26-01-10
- Mo, Y. (2008). Duration and intensity as perceptual cues for naïve listeners' prominence and boundary perception. *Proceedings of Speech Prosody 2008*, 739–742.
- Montrul, S., and Polinsky, M. (2011). Why not heritage speakers? *Linguistics. Approach. Bilingual.* 1, 58–62. doi: 10.1075/lab.1.1.07mon

- Neeleman, A., and Titov, E. (2009). Focus, contrast, and stress in Russian. *Linguistic Inquiry* 40, 514–524. doi: 10.1162/ling.2009.40.3.514
- Ortega-Llebaria, M., and Colantoni, L. (2014). L2 English intonation: relations between form-meaning associations, access to meaning, and L1 transfer. *Stud. Second. Lang. Acquis.* 36, 331–353. doi: 10.1017/S0272263114000011
- Patil, U., Kentner, G., Gollrad, A., Kügler, F., Féry, C., and Vasishth, S. (2008). Focus, word order, and intonation in Hindi. *J. South Asian Linguis.* 1, 53–70.
- Pierrehumbert, J., and Hirschberg, J. B. (1990). “The meaning of intonational contours in the interpretation of discourse” in *Intentions in communication*. eds. P. R. Cohen et al., vol. 1990 (Cambridge, MA: MIT Press), 271–311.
- Sekerina, I. A. (1999). The scrambling complexity hypothesis and processing of split scrambling constructions in Russian. *J. Slavic Linguis.*, 265–304.
- Selkirk, E. (1995). Sentence prosody: intonation, stress, and phrasing. *Handbook Phonol. Theory* 1, 550–569.
- Šimik, R. (in press). Prosodic reflexes of information structure. To appear in the *Cambridge handbook of Slavic linguistics*: Danko Šipka and Wayles Browne.
- Sioussar, N. (2007). Grammar and information structure. A study with reference to Russian, University of Utrecht/LOT: PhD thesis.
- Sorace, A. (2007). Optionality at the syntax-discourse interface in near-native L2 speakers. *Second Language*, 6, 3–15.
- Sorace, A. (2011). Pinning down the concept of “interface” in bilingualism. *Linguis. Approach. Bilingual.* 1, 1–33. doi: 10.1075/lab.1.1.01sor
- Sorace, A., and Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second. Lang. Res.* 22, 339–368. doi: 10.1191/0267658306sr271oa
- Sorace, A., and Serratrice, L. (2009). Internal and external interfaces in bilingual language development: beyond structural overlap. *Int. J. Biling.* 13, 195–210. doi: 10.1177/1367006909339810
- Takahashi, C., Kao, S., Baek, H., Yeung, A. H., Hwang, J., and Broselow, E. (2018). Native and non-native speaker processing and production of contrastive focus prosody. *Proceed. Linguis. Soc. Amer.* 3, 35–31. doi: 10.3765/plsa.v3i1.4323
- Titov, E. (2012). Information structure of argument order alternations. PhD Thesis, UCL (University College London).
- Titov, E. (2017). The canonical order of Russian objects. *Linguis. Inq.* 48, 427–457. doi: 10.1162/ling_a_00249
- Vainio, M., and Järviö, J. (2006). Tonal features, intensity, and word order in the perception of prominence. *J. Phon.* 34, 319–342. doi: 10.1016/j.wocn.2005.06.004
- Wagner, M., Breen, M., Flemming, E., Shattuck-Hufnagel, S., and Gibson, E. (2010). “Prosodic effects of discourse salience and association with focus,” in *Proceedings of the Conference on Speech Prosody*, (Chicago, IL, USA) 1–4. doi: 10.21437/speechprosody.2010-251
- Watson, D. G. (2010). “The many roads to prominence: understanding emphasis in conversation” in *Psychology of learning and motivation*, vol. 52 (Academic Press), 163–183.
- Wolk, L., Abdelli-Beruh, N. B., and Slavin, D. (2012). Habitual use of vocal fry in young adult female speakers. *J. Voice* 26, e111–e116. doi: 10.1016/j.jvoice.2011.04.007
- Xu, Y., and Xu, C. X. (2005). Phonetic realization of focus in English declarative intonation. *J. Phon.* 33, 159–197. doi: 10.1016/j.wocn.2004.11.001
- Yeung, A. H. L., Baek, H., Takahashi, C., Duncan, J., Benedette, S., Hwang, J., et al. (2019). Pitch range, intensity, and vocal fry in non-native and native English focus intonation. *Proceed. Meet. Acoust.* 36.
- Zybatow, G., and Mehlhorn, G. (2000). “Experimental evidence for focus structure in Russian” in *Annual workshop on formal approaches to Slavic linguistics* (Ann Arbor: Michigan Slavic Publications), 414–434.



OPEN ACCESS

EDITED BY

Stavros Skopeteas,
University of Göttingen, Germany

REVIEWED BY

Cristina Guardiano,
University of Modena and Reggio Emilia, Italy
Balkiz Ozturk,
Boğaziçi University, Türkiye

*CORRESPONDENCE

Ioanna Sitaridou
✉ is269@cam.ac.uk

RECEIVED 13 November 2023

ACCEPTED 26 February 2025

PUBLISHED 19 May 2025

CITATION

Neocleous N and Sitaridou I (2025) Word
order and information structure in Romeyka.
Front. Psychol. 16:1337962.
doi: 10.3389/fpsyg.2025.1337962

COPYRIGHT

© 2025 Neocleous and Sitaridou. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Word order and information structure in Romeyka

Nicolaos Neocleous¹ and Ioanna Sitaridou^{2*}

¹Aristotle University of Thessaloniki, Greece, United Kingdom, ²Faculty of Modern & Medieval Languages and Linguistics Raised Faculty Building University of Cambridge and Queens' College, Cambridge, United Kingdom

Introduction: This study examines the organization of information structure in Romeyka, the only surviving variety of Asia Minor Greek still spoken in present-day Anatolia, Turkey. Given its historical isolation from Modern Greek and its prolonged contact with Turkish, Romeyka presents a unique linguistic environment for analyzing the structural roles of [focus] and [topic].

Methods: Using empirical data, we investigate how [focus] and [topic] are realized in Romeyka. We analyze their structural positioning within the left periphery and examine their association with an *ex situ* realization.

Results: Our findings indicate that [focus] and [topic] function as independent structural notions in Romeyka. Both elements are consistently positioned in the left periphery, suggesting a systematic approach to information structuring distinct from Modern Greek.

Discussion: The observed patterns provide evidence of a reorganization of information structure in Romeyka, likely influenced by its long-term linguistic isolation and contact with Turkish. Comparisons with Pontic Greek highlight both similarities and differences, offering insights into the potential contact-induced changes in Romeyka's grammar.

KEYWORDS

focus, topic, word order, information structure, OV/VO alternation, headedness, Romeyka, Pontic Greek, Standard Modern Greek

1 Introduction

In this article, we investigate, in turn, two aspects of the grammar of Romeyka¹, an endangered Asia Minor Greek variety, namely, its word order and information structure. Romeyka is notable for being the only Asia Minor Greek variety still spoken in the Black Sea region of Turkey and as such has attracted significant theoretical interest in recent years, regarding infinitives and complementation, *wh*-formation, double-object constructions, negation, etc. (see Sitaridou, 2013, 2014a, 2014b, 2015, 2017a, 2017b, 2021, 2022, 2023a, 2023b, 2023c; Michelioudakis and Sitaridou, 2013, 2016, 2020, *inter alia*). However, neither Romeyka's underlying word order nor the organisation of its information structure has previously been subject to discussion in either descriptive or formal terms.

Specifically, we make two contributions: (i) we provide evidence that Romeyka has underlying head-initial word order in the verbal domain, that is, verb-object (VO), with the verb raising to T⁰; and (ii) we argue for a clausal architecture in Romeyka whereby [topic] and [focus] both constitute autonomous structural notions, realised as projections in the clausal left periphery, and hosting *ex situ* topicalised and focussed constituents, respectively. This study also has a

¹ Throughout the article, we use the term "Romeyka" to refer to the Greek variety, which is spoken in Turkey's Black Sea today, whereas we use the term "Pontic Greek" to refer to the Pontic Greek variety which is spoken in Greece today. For glossonymy, see Sitaridou (2023a, 2023b) and references therein.

micro-comparative element, as throughout we compare Romeyka with its cognate variety, namely, Pontic Greek (PG) (see [Sitaridou and Kaltsa, 2014](#)); we also compare Romeyka with Standard Modern Greek (SMG) and Turkish. This is relevant in view of Romeyka's socio-historical profile: being spoken in Turkey, Romeyka has for centuries developed in semi-isolation from other Greek varieties, in contact with Turkish instead (see [Neocleous, 2022](#); [Neocleous and Sitaridou, 2022](#); [Sitaridou, 2013, 2015](#)).

Previous research has established that Romeyka exhibits both frequent VO and OV order in matrix declarative clauses (see [Neocleous, 2020](#); [Neocleous and Sitaridou, 2022](#); [Sitaridou, 2014b](#); and also [Michelioudakis and Sitaridou, 2012, 2013, 2016, 2020](#); [Sitaridou, 2013, 2014a, 2014b, 2015, 2017a, 2017b, 2021, 2022](#), in prep; [Sitaridou and Kaltsa, 2014](#) for word order in Pontic Greek). Specifically, three permutations of subject (S), verb (V), and object (O) are found in such clauses: SVO (see (1)a), SOV (see (1)b), and OSV (see (1)c), but not V-initial and S-final word orders, namely, VSO (see (2)a), VOS (see (2)b), or OVS (see (2)c)—in contrast to SMG. Subordinate declarative clauses, on the other hand, only permit SOV and OSV when finite (see (3)a–(3)b), though they are obligatorily (S)VO when the embedded verb is non-finite (see (4)).

(1) Matrix declarative clauses:

a. SVO clause:

o dohtóris epíren tin aišé.
the.NOM doctor.NOM marry.PST.3SG the.ACC Ayşe.ACC
'The doctor married Ayşe.'
(S01; 140102_00080; 01:25)

b. SOV clause:

o dohtóris tin aišé epíren.
the.NOM doctor.NOM the.ACC Ayşe.ACC marry.PST.3SG
'The doctor married Ayşe.'
(S01; 140102_0008; 01:41)

c. OSV clause:

tin aišé o dohtóris epíren.
the.ACC Ayşe.ACC the.NOM doctor.NOM marry.PST.3SG
'The doctor married Ayşe.'
(S01; 140102_0008; 01:33)

(2) Not-attested matrix declarative clauses:

a. VSO clause:

?epíren o dohtóris tin aišé.
marry.PST.3SG the.NOM doctor.NOM the.ACC Ayşe.ACC
'The doctor married Ayşe.'

b. VOS clause:

?epíren tin aišé o dohtóris.
marry.PST.3SG the.ACC Ayşe.ACC the.NOM doctor.NOM
'The doctor married Ayşe.'

c. OVS clause:

?tin aišé epíren o dohtóris.
the.ACC Ayşe.ACC marry.PST.3SG the.NOM doctor.NOM
'The doctor married Ayşe.'

(3) Finite subordinate declarative clauses:

a. SOV clause:

o mohal:ímis ípen, i jylsén aténan utš
the.NOM teacher.NOM say.PST.3SG the.NOM Gülsen.
NOM she.ACC NEG
ayapá.
love.3SG

'The teacher said that Gülsen doesn't like her'
(S02; 812_0065; 05:06)

b. OSV clause:

o mohal:ímis ípen, aténan i jylsén utš
the.NOM teacher.NOM say.PST.3SG she.ACC the.
NOM Gülsen.NOM NEG
ayapá.
love.3SG
'The teacher said that Gülsen doesn't like her'
(S02; 812_0065; 05:01)

(4) Non-finite subordinate declarative clauses:

(S)VO clause:

na mutš íxa šíta spundžisíni t ospítin,
PRT.MOD NEG have.IPFV.1SG clean.INF the.
ACC house.ACC
n épezes me ta xómatae.
PRT.MOD play.IPFV.2SG with the.ACC soil.ACC
'If I hadn't just cleaned the house, you would have played with the soil.'
(S01; 812_0123; 03:32)

This is like Pontic Greek, in that it also exhibits variation between VO and OV orders. [Sitaridou and Kaltsa \(2014\)](#) argue that this belies an underlying order of VO, with OV orders derived from the discourse-driven fronting of objects. Specifically, they argue that Pontic Greek's information structure is organised as in (5), with a dedicated landing-site for information focus (IFocP), as well as a higher Contrast projection hosting contrastive topics and foci (ContrastP) and realised by a *pa*-particle, plus up to two dedicated topic positions (TopP).

(5) Pontic Greek:

TopicP ... ContrastP ... (TopicP) ... IFocP ... TP
([Sitaridou and Kaltsa, 2014](#): p. 23)

In what follows, we will claim that, although Romeyka's information structure has several properties in common with Pontic Greek, it nonetheless differs from both PG and SMG. On the one hand, we argue Romeyka too has underlying VO order with the verb raising to T⁰, deriving OV from discourse operations; and that, like Pontic Greek but unlike SMG, both focus and topichood are associated with *ex situ* realisations in the left periphery. However, in contrast to Pontic Greek but interestingly parallel to Turkish—the majority and major contact language, we conclude that Romeyka has only a *single* designated focus projection, regardless of the semantic type of focus (information or contrastive); as well as a *single* designated topic position, regardless of the semantic type of topic (aboutness or contrastive).

The article is structured as follows. §2 addresses methodological issues. §3 considers diagnostics for Romeyka's underlying word order; before a detailed investigation of the distribution of topics and foci, respectively, in Romeyka is presented in §4, offering comparisons with Pontic Greek, SMG, and Turkish as relevant. Drawing on this, §5 then provides a clausal architecture for Romeyka's left periphery. The article concludes in §6.

2 Methodology of data collection

Romeyka is spoken in three enclaves in Turkey's Black Sea region: Çaykara, Sürmene, and Tonya (see [Deffner, 1878](#); [Mackridge, 1987](#),

1999; Sitaridou, 2013, 2015, 2023a, 2023b, 2023c, 2024; Özkan, 2013; Parcharidis, 1880; Sağlam, 2017; Schreiber and Sitaridou, 2017).

The results reported here were obtained from three corpora consisting of data collected in a remote part of the Çaykara region; specifically, in a village which we refer to as ‘Anasta’ to preserve its anonymity and that of the informants (following Sitaridou, 2013: p. 104). In this study, we use data from two speakers: S01 and S02. Both speakers are female: S01 was in her 40s and S02 in her 70s when interviewed in 2015; the eldest has Romeyka L1 and Turkish L2, and the other one is a more balanced bilingual. In particular, these corpora comprise:

- a. 34 examples (from S01) from a corpus consisting of data collected by Dr Nicolaos Neocleous during a field trip in the village of ‘Anasta’, Çaykara, Black Sea, Turkey in July 2015, under the guidance, mentoring and supervision of Prof Ioanna Sitaridou who made all arrangements for this field trip and who was there in person for the entire duration, comprising 18 files and amounting to 02:51:43.
- b. 17 examples (from S01 and S02) from a corpus consisting of audio recordings collected during fieldwork by Prof Ioanna Sitaridou in the village of ‘Anasta’, Çaykara, Black Sea, Turkey in 2012 and 2014, comprising 43 files and amounting to 11:06:11.
- c. 1 example (from S01) from a corpus consisting of audio recordings collected by Prof Ioanna Sitaridou during a field trip in the village of ‘Anasta’, Çaykara, Black Sea, Turkey in July 2015, comprising 51 files and amounting to 08:25:14.

Throughout this study, we draw on data from all three corpora.² Importantly, the two females are the same in all corpora—first interviewed by Prof. Sitaridou in 2009—so the present corpus consistently describes the grammar of these two speakers. Finally, we also draw data and comparisons on both Romeyka and Pontic Greek—Romeyka’s closest cognate—from a body of previously published studies (see Michelioudakis and Sitaridou, 2012, 2013, 2016; Sitaridou, 2013, 2014a, 2014b, 2015, 2017a, 2017b, 2021, 2022, in prep; Sitaridou and Kaltsa, 2014; Neocleous and Sitaridou, 2022) Crucially, the same grammar has been consistently described by Sitaridou in all her works and collaborations.

The data collection involved oral interviews based on structured questionnaires as well as (semi-)spontaneous data. The data were audio recorded. The audio recordings were transcribed in the International Phonetic Alphabet (IPA) and annotated for the purposes of the study.

3 Diagnosing underlying word order: OV or VO?

We focus first on the underlying word order in the verbal domain—that is, whether Romeyka is underlyingly head-initial (VO) or head-final (OV)—before addressing verb placement. Given the availability of both VO and OV orders in matrix clauses attested above (see (1)–(3)), the question of underlying word order is clearly

TABLE 1 Subject agreement paradigm in Romeyka.

Person	Singular number	Plural number
1st	léy-o	léy-umen
2nd	léj-is	léj-ete
3rd	léj-i	léy-un(e)

non-trivial. In what follows, we apply several tests to distinguish the relevant possibilities.

A first source of evidence, which forms the basis for several tests, is the word order which surfaces in “pragmatically unmarked” contexts. These are contexts which afford no single constituent a special discourse-oriented interpretation such as topic or focus; such contexts are thus informative because, in principle, they control for discourse-driven movement operations and thus may better reflect the underived, underlying order. One test for pragmatically unmarked order involves ‘all-focus’ questions, exemplified by ‘What happened?’. Since these typically involve a context in which all the information expressed in the answer constitutes new information, no *single* constituent has a special discourse-oriented interpretation—all are equally focussed. Thus, the answer is ‘pragmatically unmarked’ (Büiring, 2009; van der Wal, 2016). Importantly, this test yields VO order as pragmatically unmarked in Romeyka (see (6)), just as it does for SMG, a language trivially analysed as underlyingly VO (see (7)).

- (6) Romeyka:
 - a. Question: do ejéndo?
what.NOM happen.PST.3SG
‘What happened?’
 - b. Answer:
o mustafás epelæpsen to xoráfin.
the.NOM Mustafas.NOM put.fertiliser.PST.3SG the.
ACC field.ACC
‘Mustafas put fertiliser on the field.’
(S01; 150703_0040; 02:16)
- (7) Modern Greek:
 - a. Question:
‘What happened?’
 - b. Answer:
épase ti lába o jánis.
break.PST.3SG the.ACC lamp.ACC the.NOM Yanis.NOM
‘Yanis broke the lamp.’
Alexiadou and Anagnostopoulou

Further evidence that Romeyka’s underived word order is VO comes from the behaviour of non-finite subordinate clauses. As above, such clauses are obligatorily VO (see (4)); that is, the matrix clause auxiliary verb must precede the non-finite embedded verb, which must precede the object of the embedded clause, yielding Aux-V-O, as in (8). Given that *only* VO is licit in this context it is plausible to assume that VO is the syntactically unmarked, i.e., basic word order; that is, Romeyka is underlyingly VO. Crucially, in Cappadocian, we find the opposite pattern, namely Aux-O-V orders (see Neocleous and Sitaridou, 2022).

2 See also Neocleous (2020) which is using the same corpora as the ones detailed here. It is duly noted that his PhD Supervisor, Prof Ioanna Sitaridou, has generously shared her corpus with him, for the exclusive purpose of writing up his thesis and the co-publication of any papers deriving from the thesis.

(8) Romeyka:

n íxame paníni χtisíni t ospít
 PRT.MOD have.PST.3PL go.INF build.INF the.
 ACC house.ACC
 so parjár!
 in.the.ACC pasture.ACC
 ‘I wish we had gone to build the house in the highland pastures!’
 (Sitaridou, 2014b: p. 136)

A second question more indirectly related to word order, is the position in which the verb surfaces. We present evidence that it surfaces in T⁰, at least in matrix declarative clauses; that is, Romeyka exhibits V⁰-to-T⁰ raising. In this way, Romeyka patterns with SMG, which is standardly held to exhibit verb-raising.

To start with, Romeyka exhibits two typological traits which at least frequently correlate with V⁰-to-T⁰ raising; this is evidence, albeit only suggestive, in favour of a raising analysis. On the one hand, Romeyka shows rich person and number agreement; this is consistent with the Rich Agreement Hypothesis (see Pollock, 1989; Roberts, 1993; Vikner, 1995, 1997), the generalisation that V⁰-to-T⁰ raising correlates with and according to the RAH is conditioned by, rich subject agreement on the finite verb. For example, Romeyka has distinct verbal forms for all persons, and singular and plural numbers with no suppletion, at least in most tense–voice combinations. Table 1 shows the present active declension of *léyo* ‘I say’.

However, the RAH has been disputed: empirically, on the grounds that even poorly inflected languages have been claimed to exhibit V⁰-to-T⁰ movement; and theoretically, because contemporary post-syntactic insertion models of morphology dilute its ability to influence syntax, and thus rule out rich morphology directly conditioning syntactic movement (though see Koenenman and Zeijlstra, 2014 for a rebuttal). Nonetheless, since a(n imperfect) correlation holds, this evidence is still suggestive. It is also informative that Romeyka has high tense synthesis, since on Biberauer and Roberts’ (2010) alternative account, it is tense synthesis which correlates with verb-raising instead; again, suggesting evidence for verb-raising in Romeyka.

The same is true of the second typological trait, pro-drop, whereby a clause contains no overt subject (see (9)).

(9) Romeyka:

opsé χars ípe tes.
 yesterday now say.PST.3SG she.ACC
 ‘Yesterday she immediately told her.’
 (S01; 0120713192027; 01:36)

Several accounts of pro-drop postulate an indirect relationship with V⁰-to-T⁰. Alexiadou and Anagnostopoulou (1998), for example, propose that verb movement to T⁰ is sufficient to identify T⁰’s formal features, voiding the requirement that a subject occupy [Spec,TP] (the Extended Projection Principle of Chomsky, 1982), so permitting pro-drop. Approaches of this sort are problematic in view of recent theoretical developments: as Holmberg (2005) points out, if verb-raising is enough to value T⁰’s phi-features, then these must be base-generated as valued either on V⁰ or T⁰—unexpected if semantically uninterpretable features are always base-generated as *unvalued* as in Chomsky (2001). Nonetheless, to the extent that there is again an empirical correlation between verb-raising and pro-drop, this is again suggestive evidence that Romeyka has verb-raising.

A more robust argument for V⁰-to-T⁰ raising comes from placement facts. Consider the additive particle *dže* ‘also’. *dž(e)* is like

Modern Greek *ce* ‘also’: it is a focal associate operator which surfaces only in a base-generated position as a sister to its associate; we take this position to be [Spec, vP], following Chatzikyriakidis et al.’ (2015) analysis of *ce*. Crucially, as example (10) shows, finite verbs precede *dž(e)*—providing strong evidence that the verb must raise out of vP, i.e., to T⁰, and thus past *dž(e)*.

(10) Romeyka:

Context:
eftámen vútiron. ta tsupađítikæ t alévraε yavurévumen. thénumen dže neron. thénumen dže álas. evrázumen to nerón.
 “We add butter, we fry the flour in the butter, we also put water, we also put salt, we boil the water.”
 thénumen dže álas.
 put.1PL PRT salt.ACC
 ‘We also put salt.’
 (S01; 150703_0041; 05:09)

Thus, we conclude that Romeyka does exhibit V⁰-to-T⁰ movement; and, on the strength of the evidence above, that it has underlying VO word order.

4 Information structure: distribution of topics and foci in Romeyka

We currently turn to the second goal of this article: to probe the organisation of information structure in Romeyka. This will explain, among other things, the existence of those other pragmatically marked, OV word orders. In this section, we consider, in turn, the distributions of topics and then foci in Romeyka, comparing this in each case with Romeyka’s cognate variety, Pontic Greek. This evidence forms the basis for our proposal for the architecture of the left periphery in Romeyka, discussed in §5.

4.1 Topics

We begin with topics, concentrating specifically on topics of two kinds: aboutness topics and contrastive topics. Aboutness topics are identified in the literature as the constituent which represents the theme of the predication, i.e., what the sentence is about (see Frascarelli and Hinterhölzl, 2007). A contrastive topic, on the other hand, is the sort of interpretation favoured for a constituent in a context where the hearer answers a question differing from the one being asked; that is, the constituent contrasts with some contextually salient alternative (see Büring, 2003, 2009).

Romeyka employs four syntactic strategies to convey topics in discourse. First, a constituent may be realised *ex situ*. Specifically, it may be left-dislocated, appearing preverbally and interpreted either as an aboutness (11) or a contrastive topic (12). This is unlike SMG, for example, which largely restricts left dislocation to contrastive constituents (see §5 below but see Gryllia, 2008 for a different view).

(11) Romeyka:

a. Question: tin aishén ts epíren?
 the.ACC Ayşe.ACC who.NOM marry.PST.3SG

‘Who married Ayşe?’

b. Answer:

[tin aišén]_{A-Top} o mohalímis epíren.
the.ACC Ayşe.ACC the.NOM teacher.NOM marry.PST.3SG
‘The teacher married Ayşe.’
(S01; 140102_0008; 01:10)

(12) Romeyka:

- a. Context: to pont^hólin alís epíren,
the.ACC trousers.ACC Alis.NOM buy.PST.3SG
to kazáçin o mehmétis epíren.
the.ACC sweater.ACC the.NOM Mehmetis.
NOM buy.PST.3SG
‘Alis bought the trousers and Mehmetis bought the sweater.’
- b. Question: to pont^hólin ts epíren
the.ACC trousers.ACC who.NOM buy.PST.3SG
tše to kazáçin ts epíren?
and the.ACC sweater.ACC who.NOM buy.PST.3SG
‘Who bought the trousers and who the sweater?’
- c. Answer:
[to pont^hólin]_{C-Top} alís epíren,
the.ACC trousers.ACC Alis.NOM buy.PST.3SG
áma [to kazáçin]_{C-Top} o mehmétis epíren.
but the.ACC sweater.ACC the.NOM Mehmetis.
NOM buy.PST.3SG
‘Alis bought the trousers, but Mehmetis bought the sweater.’
(S01; 150702_0014; 05:10)

Second, an aboutness topic—but not a contrastive topic (see (13) b)—may be yielded through clitic left dislocation (CLLD) with the only clitic attested in Romeyka (see Sitaridou, 2017b), i.e., *æ* ‘him/her/it/them’ (see (13)a):

(13) Romeyka:

- a. ombrón [ta patsiðæ]_{A-Top}
in.the.past the.ACC girls.ACC
s ok^húl:in tš epóliyan æ.
to school.ACC NEG send.IPFV.3PL them
‘In the past, they did not send the girls to school.’
(S01; 150702_0019; 03:23)
- b. ?ombrón [ta patsiðæ]_{C-Top}
in.the.past the.ACC girls.ACC,
s ok^húl:in tš epóliyan æ.
to school.ACC NEG send.IPFV.3PL them
‘In the past, they did not send the girls to school.’

Interestingly, although CLLD also occurs in SMG, it does not have the same pragmatic import as in Romeyka. While in SMG a left-dislocated constituent is interpreted as a topic if and only if it is CLLD’ed (see (14)a), otherwise being interpreted as a focus (see (14) b), in Romeyka a left-dislocated constituent (even if it is a definite DP) can be interpreted as a topic even if it is not CLLD’ed (see a CLLD’ed topic in (15)a and a non-CLLD’ed one in (15)b. CLLD’ed topics in Romeyka cannot be C-Top (see (15)c):

(14) Modern Greek:

- a. [to jáni]_{Top} ton sinádisa χθες.
the.ACC Yanis.ACC he.ACC meet.PST.1SG yesterday

‘I met Yanis yesterday.’

- b. [to jáni]_{Foc} (*ton) sinádisa χθες.
the.ACC Yanis.ACC he.ACC meet.PST.1SG yesterday
‘It is Yanis that I met yesterday.’
(Tsimpli, 1995: p. 179)

(15) Romeyka:

- a. [ta patátes]_{A-Top} zimónum æ.
the.ACC potatoes.ACC knead.1PL them
‘We knead the potatoes.’
(S01; 150702_0019; 05:52)
- b. Context:
tróyum ata. eftém æ me ta patátes. me ta jeralmasiæ. kaθarízum æ. ta jeralmasiæ kuzardévum æ.
‘We eat them. We make them with potatoes. With potatoes. We peel and slice the potatoes.’
[ta patátes]_{A-Top} zimónum.
the.ACC potatoes.ACC knead.1PL
‘We knead the potatoes.’
(S01; 150702_0019; 06:25)
- c. [ta patátes]_{C-Top} zimónum æ.
the.ACC potatoes ACC knead.1PL them
‘We knead the potatoes.’

The third strategy is the use of a topic particle, i.e., *pa(l)*. This assigns contrastive (but not aboutness—see (16)b) topichood to the constituent with which it is associated (see (16)a):

(16) Romeyka:

- a. eyó [ton p^hará pal]_{C-Top} ðíyo se.
I the.ACC money.ACC PRT give.1SG you.ACC
‘I give you the money.’
(S01; 0120713192027; 01:49)
- b. ?eyó [ton p^hará pal]_{A-Top} ðíyo se.
I the.ACC money.ACC PRT give.1SG you.ACC
‘I give you the money.’

Fourth, given information *may* appear postverbally, but only if non-contrastive (see (17)):

(17) Romeyka:

- a. Question:
Píos epíren tin aišén?
who.NOM marry.PST.3SG the.ACC Ayşe.ACC
‘Who married Ayşe?’
- b. Answer:
o dohtóris epíren tin aišén.
the.NOM doctor.NOM marry.PST.3SG the.ACC Ayşe.ACC
‘The doctor married Ayşe.’
(S01; 140102_0008; 01:25)

In sum, then, contrastive topics in Romeyka must be realised *ex situ* in preverbal position though they may also be marked with the *pa(l)* topic particle; non-contrastive given information, like aboutness topics, may also be realised *ex situ*, though can also be encoded *in situ* or with CLLD.

This is similar in important respects to Pontic Greek. Pontic Greek has *two* main strategies for conveying old information, both of which

have parallels in Romeyka: (a) CILD and (b) usage of a particle, *pa* (Sitaridou and Kaltsa, 2014: p. 6). CILD is exemplified by (18)a-b. Interestingly, Pontic Greek and Romeyka also both show the same prohibition on clitic doubling with right dislocation, not found in SMG (Sitaridou and Kaltsa, 2014: p. 6); compare Pontic Greek (18)c-d with Romeyka (19)b.

(18) Pontic Greek:

- a. tin elean eðek aten to jitonan.
the.ACC olive.ACC give.PST.1SG her the.
ACC neighbour.ACC
'I gave the olive to the neighbour.'
- b. ?ton jitonan eðek aton din elean.
the.ACC neighbour.ACC give.PST.1SG he.ACC the.
ACC olive.ACC
- c. *eðek aten to jitonan din elean.
give.PST.1SG her the.ACC neighbour.ACC the.
ACC olive.ACC
- d. *eðek aton to jitonan din elean.
give.PST.1SG he.ACC the.ACC neighbour.ACC the.
ACC olive.ACC
(Sitaridou and Kaltsa, 2014: p. 6 *apud* Drettas, 1997: p. 278)

(19) Romeyka:

- a. ta patátes zimónum æ.
the.ACC potatoes.ACC knead.1PL them
'We knead the potatoes.'
(S01: 150702_0019; 05:52)
- b. zimónum æ ta patátes. knead.
1PL them the.ACC potatoes.ACC
'We knead the potatoes.'

The *pa*-particle, on the other hand, attaches at the right edge of a (contrastively) topicalised constituent. The *pa*-marked constituent must also be realised *ex situ* in preverbal position as *tin aðelfi s pa* 'your sister' is in (20). This is again like Romeyka, which we have shown permits *ex situ* topics, as (21) again attests.

(20) Pontic Greek:

- tin aðelfi s pa m æyliyoris.
the.ACC sister.ACC you.POSS PRT NEG forget.2SG
'As for your sister, don't rush (into marrying her).'
- (Sitaridou and Kaltsa, 2014: p. 8 *apud* Melanofrydis, 2001: p. 13)

(21) Romeyka:

- eyó ton p'hará pal ðíyo se.
I the.ACC money.ACC PRT give.1SG you.ACC
'I give you the money.'
(S01: 0120713192027; 01:49)

However, Pontic Greek's *pa* and Romeyka's *pa(l)* are not exactly analogous. As discussed at length by Sitaridou and Kaltsa (2014), *pa* functions as a contrastive particle in Pontic Greek, realising the head of a ContrastP projection in the left periphery. Romeyka's *pa(l)*, conversely, does not seem to encode contrast, being more rigidly associated with topichood (instead of contrast). Rather, *pa(l)* seems to reflect the stage prior to the one we find in Pontic Greek, where

grammaticalisation of contrast into the particle has not occurred; Pharsiot and Rumeic Greek also seem to instantiate this stage (see Agouraki, 2010; Dawkins, 1916; Kisilier, 2007). Given this, we might consider Romeyka's *pa(l)*-phrases to instantiate the head of a dedicated TopP instead, rather than ContrastP.

4.2 Information foci

We currently turn to the distribution of focussed constituents, starting with information focus. Focus can be defined as the part of the sentence which is not presupposed (Jackendoff, 1972; Chomsky, 1972). The information focus constitutes the assertion of an utterance, i.e., its non-presupposed content, without any further restrictions; it simply asserts the membership of an individual in a set (see Gundel, 1998).

The most widely accepted test for focussed constituents is to use *wh*-questions and their answers (Beaver and Clark, 2008; Krifka, 2007; Lambrecht, 1994; Rooth, 1992; van der Wal, 2016, i.a.). A *wh*-question always yields new information, relating to the *wh*-questioned constituent; accordingly, if focus is defined as the new (i.e., non-presupposed) information in a sentence, then it follows that the phrase that replaces the *wh*-constituent is focussed.

We apply this test to Romeyka in (22) and (23). These examples clearly demonstrate that Romeyka forces information-focussed constituents to appear preverbally; the focussed objects *çavíts* 'pudding' and *pol:á chitápæ* 'many books' both appear *ex situ*, immediately left adjacent to the verb. Indeed, the infelicity of (22)c and (23)c, with the focussed constituents in postverbal position, suggests that the preverbal realisation is obligatory.

(22) Romeyka:

- a. Question:
alís dóyna éfaen?
Alis.NOM what.ACC eat.PST.3SG
'What did Alis eat?'
- Answers:
- b. alís [çavíts]_{I-Foc} éfaen.
Alis.NOM pudding.ACC eat.PST.3SG
'Alis ate a pudding.'
(S01: 150703_0040; 07:14)
- c. #alís éfaen [çavíts]_{I-Foc}.
Alis.NOM eat.PST.3SG pudding.ACC
'Alis ate a pudding.'

(23) Romeyka:

- a. Question:
dó exújepsen?
what.ACC read.PST.3SG
'What did s/he read?'
- Answers:
- b. [pol:á c'hiTÁpæ]_{I-Foc} exújepsen.
many.ACC books.ACC read.PST.3SG
'S/He read many books.'
(S01: 812_0059; 00:10)
- c. #exújepsen [pol:á c'hiTÁpæ]_{I-Foc}.
read.PST.3SG many.ACC books.ACC
'S/He read many books.'

This is unlike SMG, in which information focus has traditionally been said to occur only postverbally, as (24) and (25) suggest.

(24) Modern Greek:

- a. Question: *ti éfaje o jóryos?*
 what.ACC eat.PST.3SG the.NOM George.NOM
 ‘What did George eat?’
- b. Answer:
o jóryos éfaje [tin kobósta]_{I-Foc}
 the.NOM George.NOM eat.PST.3SG the.ACC stewed-
 fruit.ACC
 ‘George ate the stewed fruits.’
 (Sitaridou and Kaltsa, 2014: p. 12)

(25) Modern Greek:

- a. Question: *ti ðjávase?*
 what.ACC read.PST.3SG
 ‘What did s/he read?’
- Answers:
- b. *ðjávase [polá vivlíá]_{I-Foc}*
 read.PST.3SG many.ACC books.ACC
 ‘S/He read many books.’
- c. *?[polá vivlíá]_{I-Foc} ðjávase.*
 many.ACC books.ACC read.PST.3SG
 ‘S/He read many books.’
 (Sitaridou and Kaltsa, 2014: p. 12)

This difference attenuates if we consider Gryllia’s (2008) findings, which show—based on experimental tests—that preverbal objects are neither necessarily exhaustive nor exclusively contrastive in SMG; such that information focus can be preverbal. For example, the focussed direct object is interpreted as a new information focus both when occurring in OV order, and in VO order in (26)c and (26)b, respectively.

(26) Modern Greek:

- a. Question:
ti xárise metaksí álon
 what.ACC give.PST.3SG among others.GEN
o jánis stin ilektra?
 the.NOM Yanis.NOM to.the.ACC Ilektra.ACC
 ‘Among other things, what did Yanis give to Ilektra?’
- Answers:
- b. *xárise [éna vivlíó]_{I-Foc} stin iléktra.*
 give.PST.3SG a.ACC book.ACC to.the.ACC Ilektra.ACC
 ‘He gave a book to Ilektra.’
- c. *[éna vivlíó]_{I-Foc} xárise stin iléktra.*
 a.ACC book.ACC give.PST.3SG to.
 the.ACC Ilektra.ACC
 ‘He gave a book to Ilektra.’
 (Gryllia, 2008: p. 21)

Nevertheless, the fact that SMG may in fact allow *both* options does not alter the parametric difference with Romeyka, where the preverbal position is the only option. This is particularly clear in the following judgement made by a Romeyka speaker in (27).

(27) Romeyka:

- a. Question:

- ánda erotó se alís dóyna éfaen*
 if ask.1SG you.ACC Alis.NOM what.ACC eat.PST.3SG
esí léjis me, o alís éfaen míla
 you.NOM say.2SG I.ACC the.NOM Alis.NOM eat.
 PST.3SG apples.ACC
jókxa, alís míla éfaen?
 or Alis.NOM apples.ACC eat.PST.3SG
 ‘If I ask you, what did Alis eat, what do you say to me? Alis ate apples, or Alis apples ate?’
- b. Answer:
kal:íon, alís [míla]_{I-Foc} éfaen.
 better Alis.NOM apples.ACC eat.PST.3SG
 ‘Alis ate apples, sounds better.’
 (S01; 812_0055; 03:09)

By contrast, Pontic Greek does pattern with Romeyka: information-focussed constituents in Pontic Greek also occur preverbally as demonstrated by *to xošaf* ‘the stewed fruit’ in (28) and *pola vivlia* ‘many books’ in (29) below.

(28) Pontic Greek:

- a. Question:
o jorikas do efaen?
 the.NOM Yorikas.NOM what.ACC eat.PST.3SG?
 ‘What did Yorikas eat?’
- Answers:
- b. *(o jorikas) [to xošaf]_{I-Foc} efaen.*
 the.NOM Yorikas.NOM the.ACC stewed.fruit.
 ACC eat.PST.3SG
 ‘Yorikas ate stewed fruit.’
- c. *#o jorikas efaen [do xošaf]_{I-Foc}*
 the.NOM Yorikas.NOM eat.PST.3SG the.ACC stewed.
 fruit.ACC
 ‘Yorikas ate stewed fruit.’
 (Sitaridou and Kaltsa, 2014: p. 12)

(29) Pontic Greek:

- a. Question: *do eðevasen?*
 what read.PST.3SG?
 ‘What did he read?’
- Answers:
- b. *pola vivlia eðevasen.*
 many.ACC books.ACC read.PST.3SG
 ‘He read many books.’
- c. *#eðevasen pola vivlia.*
 read.PST.3SG many.ACC books.ACC
 ‘He read many books.’
 (Sitaridou and Kaltsa, 2014: p. 12)

This suggests information focus in Romeyka may be consistent with the conclusion reached in recent research on its Pontic Greek counterpart, namely, that it appears in the left periphery (see Sitaridou and Kaltsa, 2014). For example, in both Romeyka and Pontic Greek any focussed phrase—no matter the phrase type—appears before the verb: direct object (NP) (see (30) from Romeyka and (31) from Pontic Greek), direct object (DP) (see (32) from Romeyka and (33) from Pontic Greek), indirect object (beneficiary) (DP) (see (34) from Romeyka and (35) from Pontic Greek), predicative (adjective) (see (36) from Romeyka and (37) from Pontic

Greek), adverbial (NP) (see (38) from Romeyka and (39) from Pontic Greek), and existential constructions (see (40) from Romeyka and (41) from Pontic Greek):

(30) Romeyka:

Direct object (NP) is focussed:

a. Question:

alís dóyna éfaen?

Alis.NOM what.ACC eat.PST.3SG

‘What did Alis eat?’

b. Answer:

alís [xavítsin]_{I-Foc} éfaen.

Alis.NOM pudding.ACC eat.PST.3SG

‘Alis ate a pudding.’

(S01; 150703_0040; 07:14)

(31) Pontic Greek:

Direct object (NP) is focussed:

a. Question:

do efaes?

what.ACC eat.PST.2SG

‘What did you eat?’

b. Answer:

[xavíts]_{I-Foc} efaa.

pudding.ACC eat.PST.1SG

‘I ate pudding.’

(Sitaridou and Kaltsa, 2014; p. 13 apud Drettas, 1997; p. 280)

(32) Romeyka:

Direct object (DP) is focussed:

a. Question:

i aisé tínan epíren?

the.NOM Ayşe.NOM who.ACC marry.PST.3SG

‘Who did Ayşe marry?’

b. Answer:

i aisé [ton dohtórin]_{I-Foc} epíren.

the.NOM Ayşe.NOM the.ACC doctor.ACC marry.PST.3SG

‘Ayşe married the doctor.’

(S01; 140102_0008; 01:37)

(33) Pontic Greek:

Direct object (DP) is focussed:

a. Question:

do eplises?

what.ACC wash.PST.2SG

‘What did you wash?’

b. Answer:

[ta poðaræ m]_{I-Foc} eplisa.

the.ACC feet.ACC I.POSS wash.PST.1SG

‘I washed my feet.’

(Sitaridou and Kaltsa, 2014; p. 13 apud Drettas, 1997; p. 280)

(34) Romeyka:

Indirect object (beneficiary) (DP) is focussed:

a. Question:

to c^hitápin tínan éndžes?

the.ACC book.ACC who.ACC bring.PST.2SG

‘To whom did you give the book?’

b. Answer:

to c^hitápin [ton juSÚfin]_{I-Foc} énga.

the.ACC book.ACC the.ACC Yusufis.ACC bring.PST.1SG

‘I brought the book for Yusufis.’

(S01; 150703_0042; 00:54)

(35) Pontic Greek:

Indirect object (beneficiary) (DP) is focussed:

epita ti nifæn θa eniyane lutron.

then the.ACC bride.ACC PRT.FUT open.

IPFV.3PL bath.ACC

‘Then they would prepare the bath for the married girl.’

(Sitaridou and Kaltsa, 2014; p. 13 apud Drettas, 1997; p. 280)

(36) Romeyka:

Predicative (adjective) is focussed:

a. Question:

alís do en?

Alis.NOM what.NOM be.3SG

‘What is Alis?’

b. Answer:

alís [áyuros]_{I-Foc} en.

Alis.NOM boy.NOM be.3SG

‘Alis is a boy.’

(S01; 140102_0009; 00:20)

(37) Pontic Greek:

Predicative (adjective) is focussed:

a. Question: do en atos?

what.ACC be.3SG he.NOM

‘What is he like?’

b. Answer:

palalos en.

crazy.NOM be.3SG

‘He is crazy.’

(Sitaridou and Kaltsa, 2014; p. 14 apud Drettas, 1997; p. 555)

(38) Romeyka:

Adverbial (NP) is focussed:

a. Question:

i mána s póte efáise ton musafirín?

the.NOM mother.NOM you.POSS when feed.PST.3SG the.

ACC guest.ACC

‘When did your mother feed the guest?’

b. Answer:

[opsé]_{I-Foc} efáisen ton musafirín.

yesterday feed.PST.3SG the.ACC guest.ACC

‘She fed the guest yesterday.’

(S01; 150703_0041; 07:10)

(39) Pontic Greek:

Adverbial (NP) is focussed:

mesaniyts eton.

midnight be.PST.3SG

‘It was midnight.’

(Sitaridou and Kaltsa, 2014; p. 14 apud Drettas, 1997; p. 555)

(40) Romeyka:

Existential construction is focussed:

a. Question:

o š^hil:on do en?
the.NOM dog.NOM what.NOM be.3SG
‘What is the dog?’

b. Answer:

[haivánin]_{I-Foc} en.
animal.NOM be.3SG
‘It’s an animal.’
(S01; 140102_0009; 00:35)

(41) Pontic Greek:

Existential construction is focussed:

χorafæ c^h ine.
fields NEG exist.3PL
‘There are no fields.’

(Sitaridou and Kaltsa, 2014: p. 14)

Second, the two varieties pattern together in having focus-fronting in questions of “total ignorance” that yield a yes/no reply (Sitaridou and Kaltsa, 2014: p. 14). See (42) from Romeyka and (43) from Pontic Greek:

(42) Romeyka:

esís [ta tsupáðæ]_{I-Foc} θerízete?
you.NOM the.ACC corn.ACC harvest.2PL
‘Do you harvest the corn?’
(S02; 812_0067; 01:58)

(43) Pontic Greek:

a. Question:
t apiðæ ekserts?
the.ACC pears.ACC know.2SG
‘Do you know the pears?’
b. ???ekserts t apiðæ? know.2SG the.ACC pears.ACC
‘Do you know the pears?’
(Sitaridou and Kaltsa, 2014: p. 14)

Third, just like Romeyka, Pontic Greek requires strict adjacency between the fronted information-focussed constituent and the predicate it precedes, especially where the predicate is the verb *be* or *have* (Sitaridou and Kaltsa, 2014: p. 14). In (44), for example, the adverb *panda* ‘always’ cannot interpolate between the information focussed *aiksa* ‘like this’ and the verb *esne* ‘were’, otherwise infelicity ensues (44)b.

(44) Pontic Greek:

a. aiksa esne panda.
like.this be.IP.FV.2SG always
‘You were always like this.’
b. *aiksa panda esne.
like.this always be.IP.FV.2SG
‘You were always like this.’
(Sitaridou and Kaltsa, 2014: p. 14 apud Drettas, 1997: p. 182)

Fourth, both varieties permit movement of the focussed constituent in subordinate clauses as in (45) in Romeyka and (46) in Pontic Greek (Sitaridou and Kaltsa, 2014: p. 14). Thus, the behaviour of information focus in Romeyka and Pontic Greek is highly consistent.

(45) Romeyka:

a. Question:

do θarís, alís tían efilisen?
what.ACC think.2SG Alis.NOM who.ACC kiss.PST.3SG
‘Who do you think that Alis kissed?’

b. Answer:

eyó θaró, alís [tin aišén]_{I-Foc} efilisen.
I think.1SG Alis.NOM the.ACC Ayşe.ACC kiss.PST.3SG
‘I think that Alis kissed Ayşe.’
(S01; 150703_0040; 19:07)

(46) Pontic Greek:

a. eθaresen oti tšantarmas eton.
think.PST.3SG that policeman be.IP.FV.3SG
‘He thought (that) he was a policeman.’
b. eθaresen džantarmas eton.
think.PST.3SG policeman be.IP.FV.3SG
‘He thought he was a policeman.’
(Sitaridou and Kaltsa, 2014: p. 15 apud Drettas, 1997: p. 370)

4.3 Contrastive foci

Consider finally *contrastive* focus. This involves the selection of a subset from a set of alternatives, *contrasting* with a contextually salient individual (see Molnár, 2006).

In Romeyka, contrastive focus patterns like information focus: the focussed constituent occurs preverbally as in (47). Indeed, as the Romeyka speaker’s grammaticality judgement in (48) suggests, this is the *only* option for encoding contrastive focus; again, this is like information focus.

(47) Romeyka:

a. Question:
kahVÉN jóksa tšáin thélis?
coffee.ACC or tea.ACC want.2SG
‘Do you want coffee or tea?’

Answers:

b. eyó [kahvén]_{C-Foc} thélo.
I.NOM coffee.ACC want.1SG
‘I want coffee.’
(S01; 150702_0013; 12:15)
c. manaxón [kahvén]_{C-Foc} thelo.
only coffee.ACC want.1SG
‘I only want coffee.’
(S01; 150702_0013; 12:22)

(48) Romeyka:

a. Question:
eyó léyo se alís aphíðæ ayórasen,
I.NOM say.1SG you.ACC Alis.NOM
pears.ACC buy.PST.3SG
áma esí eksérts alís míla ayórasen.
but you.NOM know.2SG Alis.NOM
apples.ACC buy.PST.3SG
eyó érxome léyo se alís aphíðæ ayórasen.
I.NOM come.1SG say.1SG you.ACC Alis.NOM
pears.ACC buy.PST.3SG
esí dóyna léjis me?
you.NOM what.ACC say.2SG I.ACC?

‘Alis bought pears, but you know that he bought apples. I came and told you that Alis bought pears. What do you reply to me?’

b. Answer:

alís [míla]_{C-Foc} ayórasen.

Alis.NOM apples.ACC buy.PST.3SG

‘Alis bought apples.’

(S01; 812_0055; 01:54)

Any type of phrase can be contrastively focussed in Romeyka, just as it can be information focussed: object (NP) (see (49)), object (DP) (see (50)), predicative complement (see (51)), adverbial phrase (see (52)), among others (see [Neocleous, 2020](#): p. 160ff):

(49) Romeyka:

Direct object (NP) is focussed:

a. Question:

o mehmétis míla jóksa aP^Hðæ ayórasen?

the.NOM Mehmetis.NOM apples.ACC or

pears.ACC buy.PST.3SG

‘Did Mehmetis buy apples or pears?’

b. Answer:

o mehmétis [míla]_{C-Foc} ayórasen.

the.NOM Mehmetis.NOM apples.ACC buy.PST.3SG

‘It’s apples that Mehmetis bought.’

(S01; 150702_0013; 12:05)

(50) Romeyka:

Direct object (DP) is focussed:

a. Question:

o ramazánis ti zeiNÉP epíren?

the.NOM Ramazanis.NOM the.ACC

Zeynep.ACC marry.PST.3SG

‘Did Ramazanis marry Zeynep?’

b. Answer:

o ramazánis [tin aiséñ]_{C-Foc} epíren.

the.NOM Ramazanis.NOM the.ACC

Ayşe.ACC marry.PST.3SG

‘It’s Ayşe that Ramazanis married.’

(S01; 140102_0009; 07:50)

(51) Romeyka:

Predicative complement is focussed:

a. Question:

dóyna en avúto? vútiron?

what.NOM be.3SG this.NOM butter.NOM

‘What is this? Butter?’

b. Answer(s):

[anθóýalan]_{C-Foc} en.

buttermilk.ACC be.3SG

‘This is buttermilk.’

(S01; 812_0055; 00:54)

(52) Romeyka:

Adverbial phrase is focussed:

a. Question:

alís osím:eron érθen asin tsáikaran?

Alis.NOM today come.PST.3SG from.the.ACC Çaykara.ACC

‘Did Alis come from Çaykara today?’

b. Answer:

jok^h, [opsé]_{C-Foc} érθen.

no yesterday come.PST.3SG

‘No, he came yesterday.’

(S01; 150703_0040; 08:46)

Unlike in the case of information focus, SMG typically realises contrastive focus by left dislocation to a preverbal position —so patterns with Romeyka. This is exemplified by (53). This is also true of Pontic Greek, as in (54).

(53) Modern Greek:

a. Question:

thélis kaFÉ i TSÁI?

want.2SG coffee.ACC or tea.ACC

‘Do you want coffee or tea?’

Answers:

b. [kaFÉ]_{C-Foc} thélo.

coffee.ACC want.1SG

‘I want coffee.’

c. móno [kaFÉ]_{C-Foc} thélo.

only coffee.ACC want.1SG

‘I only want coffee.’

([Sitaridou and Kaltsa, 2014](#): p. 12)

(54) Pontic Greek:

a. Question:

thelts na pseno se gaiven

want.2SG PRT.MOD make.1SG you.ACC coffee.ACC

ci ena ðio otia na vukuse?

and one.ACC two.ACC sweets.ACC PRT.MOD dunk.PNP.2SG

‘Do you want me to make you some coffee and a couple of sweets to dunk in the coffee?’

Answers:

b. kaiven pseson.

coffee.ACC make.IMP.2SG

‘Make coffee (and not something else).’

b’. manaxon kaiven pseson.

only coffee.ACC make.IMP.2SG

‘Only make coffee.’

c. *manaxon kaiven pa pseson.

only coffee.ACC PRT make.IMP.2SG

‘Only make coffee.’

d. kaiven pa thelo.

coffee.ACC PRT want.1SG

‘I want coffee.’

e. kaiven pa thelo, otia pa thelo.

coffee.ACC PRT want.1SG sweets.ACC PRT want.1SG

‘I want both coffee and cookies.’

([Sitaridou and Kaltsa, 2014](#): pp. 11–12)

Contrastive focus in Pontic Greek nonetheless differs from Romeyka in at least two regards. First, though it can encode contrastive focus by focus movement to preverbal position, the contrastive focussed constituent need not be strictly adjacent to the predicate. Thus, in (55) below, the left periphery elements (topicalised *aika emorfa pedja* ‘such beautiful children’, focalised *esis* ‘you’) are separated from the verb

by the adverb *kamian* ‘never,’ violating strict adjacency (Sitaridou and Kaltsa, 2014: p. 15). This contrasts with Romeyka, in which contrastive focussed constituents are obligatorily strictly adjacent to the predicate. This is parallel to the behaviour of information-focussed constituents, as noted above; thus, the strict adjacency requirement is general to all foci in Romeyka, unlike Pontic Greek.

- (55) Pontic Greek: aika emorfa peđja esis kamian iđeten?
such beautiful.ACC children.ACC
you.NOM ever see.PST.2PL
‘Have you ever seen such beautiful children?’
(Sitaridou and Kaltsa, 2014: p. 15 apud Drettas, 1997: p. 183)

The second difference between Romeyka and Pontic Greek is that the latter has another means of encoding contrastive focus which is not present in Romeyka: the use of discourse particles. This is exemplified by the particles *cela* and *ki*. Both assign contrastive focus to the constituent to which they attach, though differ somewhat in distribution. *Cela* is always in postposition, though never enclitic to the verb; in (56), it appears post-sententially, so contrastively focusing the whole VP. Conversely, *ki* is always enclitic to the verb, for example, contrasting the verbal constituent *eperane* = *ki* ‘they took’ with the predicate in the second main clause, *eksenkan* = *aten aso plan tin portant* ‘they forced her through the side door’ in (57) (Sitaridou and Kaltsa, 2014: p. 11).

- (56) Pontic Greek:
a. kit eceka ce ^h eleps ato cela.
lie.3SG there and NEG see.2SG it.ACC PRT
‘It is there and you don’t even see it’
b. efaen do fain atun c edoken atsen cela.
eat.PST.3SG the.ACC food.ACC their and strike.
PST.3SG them PRT
‘He ate their food and beat them as well.’
(Sitaridou and Kaltsa, 2014: p. 11 apud Drettas, 1997: p. 410)

- (57) Pontic Greek:
atos ... eperane ci ti marian eksenkan aten
he take.PST.3PL PRT the.ACC Maria.ACC take.out.
PST.3PL her.ACC
aso plan din bortan.
from.the.ACC sides.ACC the.ACC door.ACC
‘He ... they took Maria and forced her to exit through the side door.’
(Sitaridou and Kaltsa, 2014: p. 11 apud Drettas, 1997: p. 481)

Let us summarise our conclusions from this section. In Romeyka, topics and foci can both be expressed by occurring *ex situ*. In the case of foci, this is obligatory: focussed constituents in Romeyka always occupy the immediate preverbal position, instead of the pragmatically unmarked postverbal position (see §3), no matter the type of focus or syntactic category of the constituent. In the case of topics, there is also an option to occur in a postverbal position, but only for non-contrastive given information; contrastive topics must be preverbal, like foci. Contrastive topics can also occur with a *pa(l)*-particle.

This differs from SMG, which allows non-contrastive information—for example, aboutness topics and information foci—to occur postverbally, and does not exhibit the strict adjacency requirement on preverbal topics/foci. Romeyka also lacks topic particles like *pa(l)*. It is

strikingly more like Pontic Greek: like Romeyka, it allows foci and topics of all types to be realised preverbally; and it can also mark contrastive topics by a topic particle, *pa*. There are still differences, however: left dislocation is not obligatory for contrastive foci in Pontic Greek, for example, because focus can instead be marked by a particle; and the *pa*-particle is unlike Romeyka’s *pa(l)* in encoding contrast.

4.4 Wh-questions and focus

MG displays *wh*-questions, (see (58)). Similarly, Romeyka also employs *wh*-questions (see (59)) (see Michelioudakis and Sitaridou, 2013, 2016):

- (58) Modern Greek:
a. pços filise ti maria?
who.NOM kiss.PST.3SG the.ACC Maria.ACC
‘Who kissed Maria?’
(Alexopoulou and Baltazani, 2012)
b. pçon filise i maria?
whoACC kiss.PST.3SG the.NOM Maria.NOM
‘Who did Maria kiss?’
(Alexopoulou and Baltazani, 2012)
(59) Romeyka:
a. Płos eđótšen tin kos:áran?
who.NOM give.PST.3SG the.ACC hen.ACC
‘Who gave the hen?’
(S01; 812_0093; 00:03)
b. çavitsæ Płos éfaen?
pudding.ACC who.NOM eat.PST.3SG
‘Who ate puddings?’
(S01; 812_0057; 04:06)
c. alís DÓyna đótšen?
Alis.NOM what.ACC give.PST.3SG
‘What did Alis give?’
(S01; 812_0093; 00:16)

Crucially, the order of *wh*-questions is strictly order-preserving in Romeyka (see (60) and (61)) (see Michelioudakis and Sitaridou, 2013, 2016):

- (60) Romeyka:
a. Płos eđótšen tin kos:áran?
who.NOM give.PST.3SG the.ACC hen.ACC
‘Who gave the hen?’
(S01; 812_0093; 00:03)
b. ?Płos tin kos:áran eđótšen?
who.NOM the.ACC hen.ACC give.PST.3SG
‘Who gave the hen?’
(61) Romeyka:
a. çavitsæ Płos éfaen?
puddings.ACC who.NOM eat.PST.3SG
‘Who ate puddings?’
(S01; 812_0057; 04:06)
b. ?Płos çavitsæ éfaen?

who.NOM puddings.ACC eat.PST.3SG
 ‘Who ate puddings?’

wh-phrases are obligatorily left-dislocated (see (62)), with no option but to leave any *wh*-phrase *in situ* (Michelioudakis and Sitaridou, 2013, 2015):

- (62) Romeyka:
 a. alís DÓyna ðótšen?
 Alis.NOM what.ACC give.PST.3SG
 ‘What did Alis give?’
 (S01; 812_0093; 00:16)
 b. ?alís ðótšen DÓyna?
 Alis.NOM give.PST.3SG what.ACC
 ‘What did Alis give?’

In this section, we have shown that *wh*-phrases in *wh*-questions in Romeyka occupy the same position that focussed constituents occupy.

5 Information structure: clausal architecture of the left periphery

Having established the distributions for topics and foci, we currently move to consider what clausal structure is required to model Romeyka's topics and foci, adopting a cartographic perspective on information structure (cf. Neocleous, 2020: ch. 5 for a minimalist alternative). First, in §5.1, we consider the *number* of topic positions required in view of the data discussed in §4. We then map these positions onto a functional hierarchy for the clausal left periphery in §5.2, contrasting our proposal with the information structure systems of Pontic Greek and Turkish, respectively.

5.1 How many topic positions are there in Romeyka?

Given the evidence above, it is clear there are two positions in which given information, i.e., topics, may occur in Romeyka: a preverbal position; and a postverbal position. Note that this raises the following question: if *both* the preverbal and the postverbal domain can accommodate given information, then what differentiates these interpretatively? As we have already shown, contrastive given information can only ever appear in the preverbal domain—it is infelicitous in the postverbal domain. Thus, what differentiates the preverbal from the postverbal topics is the [contrast] feature; the preverbal, but not postverbal, encodes [contrast] to some extent at least.

The examples in (63) and (64) provide additional evidence to this effect: in (63), the object *tin aišén* ‘Ayşe’ carries [non-contrastive] given information and can occur felicitously in both preverbal (63)a and postverbal (63)b position, whereas the object *dolmán* ‘dolma’ in (64) can only appear preverbally (64)a, but not postverbally (64)b, by virtue of carrying [contrastive] given information.

- (63) Romeyka:
 a. [tin aišén]_{A-Top} o dohtóris epíren.
 the.ACC Ayşe.ACC the.NOM doctor.NOM marry.PST.3SG

‘The doctor married Ayşe.’
 b. o dohtÓris epíren [tin aišén]_{A-Top}.
 the.NOM doctor.NOM marry.PST.3SG the.ACC Ayşe.ACC
 ‘The doctor married Ayşe.’
 (S01; 140102_0008; 01:15)

- (64) Romeyka:
 a. [dolmán]_{C-Top} o mehmétis éfaen.
 dolma.ACC the.NOM Mehmetis.NOM eat.PST.3SG
 ‘Mehmetis ate dolma.’
 (S01; 150702_0014; 11:46)
 b. #o mehmÉtis éfaen [dolmán]_{C-Top}.
 the.NOM Mehmetis.NOM eat.PST.3SG dolma.ACC
 ‘Mehmetis ate dolma.’

It is important to note that we can have multiple TopP in Romeyka (see (65)):

- (65) Romeyka:
 [eyó]_{A-Top} [ton p^hará pal]_{C-Top} ðíyo se.
 I the.ACC money.ACC PRT give.1SG you.ACC
 ‘I give you the money.’
 (S01; 0120713192027; 01:49)

5.2 Clausal architecture of the left periphery

We are now able to propose an architecture for the left periphery of the Romeyka clause. Adopting a cartographic perspective, we take the focus and topic positions identified above to be realised by projections in the functional structure of the left periphery. It is worth noting at this stage that we take the external argument to raise to a high left-peripheral position in Romeyka, namely, the specifier of a (potentially iterated) TopP (for arguments to this effect, see Neocleous, 2020: pp. 105–110); this explains the ability of a topicalised/focussed object to target the left periphery but still follow the subject.

To determine how the relevant topic/focus positions, i.e., projections, are arranged hierarchically, we apply the tests used by Neeleman and van de Koot (2008) and Şener (2010) in their investigations of the information structure of Dutch and Turkish, respectively.

Consider first example (66). The context in (66)a favours an interpretation of the subject in (66)b–c, *o mehmétis* ‘Mehmetis’, as a contrastive topic, as it is the constituent which forms the expected answer. On the other hand, the object in (66)b–c, *dolmán* ‘dolma’, is interpreted as a contrastive focus. This follows from the well-known observation that, in answers to *wh*-questions, the constituent corresponding to the *wh*-operator is typically focussed (e.g., Neocleous, 2020: p. 114, Michelioudakis and Sitaridou, 2013, 2016).

- (66) Romeyka:
 a. Question:
 alís do epítšen?
 Alis.NOM what.ACC do.PST.3SG
 do éfaen so bairámin?
 what.ACC eat.PST.3SG in.the.ACC Bayram.ACC
 ‘What did Alis do? What did he eat at Bayram?’
 Answers:

- b. vál:ahi, utš ekséro alís do epítšen, áma ...
frankly NEG know.1SG Alis.NOM what.ACC do.
PST.3SG but ...
'Frankly, I don't know about Alis, but ...'
- c. [o mehmétis]_{C-Top} [dolmán]_{C-Foc} éfaen.
the.NOM Mehmetis.NOM dolma.ACC eat.PST.3SG
'Mehmetis ate dolma.'
(S01; 150702_0014; 09:06)
- d. #[dolmán]_{C-Foc} [o mehmétis]_{C-Top} éfaen.
dolma.ACC the.NOM Mehmetis.NOM eat.PST.3SG
'Mehmetis ate dolma.'

Importantly, there is a contrast between the felicitous (66)b in which the contrastive focussed constituent (C-Foc) *follows* the contrastive topic (C-Top), and the infelicitous (66)c in which C-Foc *precedes* C-Top. In other words, C-Top > C-Foc order is felicitous; C-Foc > C-Top is not.

This restriction holds even when we reverse the relation between grammatical function and information structure. The context in (67) is set up to favour an interpretation of the *object* as C-Top and the *subject* as C-Foc, the opposite of (66).

(67) Romeyka:

- a. Question:
o tšorbás do ejéndo?
the.NOM soup.NOM what.ACC happen.PST.3SG
atón kanís éfaen æ?
this.ACC anyone.NOM eat.PST.3SG it.ACC
'What about the soup? Has anyone eaten it?'
- Answers:
b. vál:ahi, utš eksér o tšorbás do ejéndo,
frankly NEG know.1SG the.NOM soup.NOM what.
ACC happen.PST.3SG
áma ...
but ...
'Frankly, I don't know about the soup, but ...'
- c. [dolmán]_{C-Top} [o mehmÉtis]_{C-Foc} éfaen.
dolma.ACC the.NOM Mehmetis.NOM eat.PST.3SG
'Mehmetis ate dolma.'
(S01; 150702_0014; 11:46)
- d. #[o mehmÉtis]_{C-Foc} [dolmán]_{C-Top} éfaen.
the.NOM Mehmetis.NOM dolma.ACC eat.PST.3SG
'Mehmetis ate dolma.'

Nonetheless, C-Top > C-Foc is again the only felicitous order as in (67)b, with C-Foc > C-Top in (67)c being infelicitous.

Third and finally, this same restriction holds for the interaction between VP-internal objects, too. The sentences in example (68) contain a ditransitive verb, where the context is set up to favour the interpretation of the IO as C-Foc and the DO as C-Top. Moreover, again, only the C-Top > C-Foc order (that is, DO > IO) order in (68) b is felicitous; C-Foc > C-Top (IO > DO) is not (68)c.

(68) Romeyka:

- a. Question:
i antíka tše i sandália do

the.NOM antique.NOM and the.NOM chair.
NOM what.ACC
ejéndo? o páp^hos tíanan éðocen æ?
happen.PST.3SG the.NOM grandfather.NOM who.
ACC give.PST.3SG it.ACC
'What about the antique table and the chair? Who did your granddad bequeath them to?'

Answers:

- b. vál:ahi i antíka do ejéndo
frankly the.NOM antique.NOM what.ACC happen.PST.3SG
utš ekséro, áma ...
NEG know.1SG but ...
'Frankly, I don't know about the antique table, but ...'
- c. [ti sandalían]_{C-Top} [ton TšÍri m]_{C-Foc} eðótšen.
the chair.ACC the father.ACC I.POSS give.PST.3SG
'my granddad bequeathed the chair to my dad.'
(S01; 150702_0014; 14:01)
- d. #[ton TšÍri m]_{C-Foc} [ti sandalían]_{C-Top} eðótšen.
the father.ACC I.POSS the chair.ACC give.PST.3SG
'my granddad bequeathed the chair to my dad.'

Moreover, in (69), where we reverse the mapping of grammatical function to information structure, such that the IO is currently interpreted as C-Top, and the DP as C-Foc, the ordering restriction still holds. This is true even though it reverses the felicity contrast as it relates to DO and IO compared to (68): currently, IO > DO is the only felicitous order, with DO > IO illicit—the opposite of (68).

(69) Romeyka:

- a. Question:
o tširis DO ejéndo?
the.NOM father.NOM what.ACC happen.PST.3SG
o páp^hos DO éðocen aton?
the.NOM grandfather.NOM what.ACC
give.PST.3SG he.ACC
'What about your dad? What has granddad bequeathed to him?'
- Answers:
a. vál:ahi o tširis m DO ejéndo
frankly the.NOM father.NOM I.POSS what.
ACC happen.PST.3SG
utš ekséro, áma ...
NEG know.1SG but
'Frankly, I don't know about my dad, but ...'
- b. [ti mána m]_{C-Top} [to saÁt^hin]_{C-Foc} efitšen.
the mother.ACC I.POSS the watch.ACC bequeath.PST.3SG
'my granddad bequeathed the watch to my mother.'
(S01; 150702_0023; 08:18)
- c. #[to saÁt^hin]_{C-Foc} [ti mána m]_{C-Top} efitšen.
the watch.ACC the mother.ACC I.POSS bequeath.PST.3SG
'my granddad bequeathed the watch to my mother.'

Thus, there is robust evidence for the generalisation that C-Top precedes C-Foc in Romeyka, independent of the grammatical functions the relevant constituents bear.

Indeed, the generalisation can be broadened to range over *all* foci. We have already established that information focus occupies the same position as contrastive focus. Accordingly, I-Foc follows contrastive topics just like C-Foc does (see (70)); and follows aboutness topics (see (71)). This strongly suggests a clausal architecture for Romeyka whereby the single dedicated Focus projection follows the (C-)Topic projection, Top > Foc.

(70) Romeyka:

- a. Question: tsi birjýlis t aðélfæ
the.GEN Birgül.GEN the.NOM brothers.NOM
d epíkane so pártin?
what.ACC do.PST.3PL at.the.ACC party.ACC
'What did Birgül's brothers get to drink at the party?'

Answers:

- a. vál:ahi as aðélfæ tes utš ekséro, áma ...
frankly from.the.ACC brothers.ACC she.
POSS NEG know.1SG but ...
'Frankly, I do not know about all her brothers, but ...'
b. [úl:unon o mikrón]_{C-Top} [raCfN]_{I-Foc} epíen.
all.GEN the.NOM young.NOM raki.ACC drink.PST.3SG
'Birgül's youngest brother drank raki.'
(S01; 150702_0023; 23:26)
c. #[raCfN]_{I-Foc} [úl:unon o mikrón]_{C-Top} epíen.
raki all.GEN the.NOM young.NOM drink.PST.3SG
'Birgül's youngest brother drank raki.'

(71) Romeyka:

- a. Question:
avúto to faín Píon patsín epítšen?
this.ACC the.ACC food.ACC which.NOM
girl.NOM make.PST.3SG
'Which girl made this food?'
(S01; 150703_0042; 03:32)
b. Answer:
[avúton to faín]_{A-Top} [i miNÉ]_{I-Foc} epítšen.
this.ACC the.ACC food.ACC the.NOM Mine.
NOM make.PST.3SG
[t ál:on]_{C-Top} [i aiŠÉ]_{C-Foc} epítšen.
the.ACC other.ACC the.NOM Ayşe.NOM make.PST.3SG
'Mine made this food; Ayşe made the other one.'
(S01; 150703_0042; 03:45)

Interestingly, this pattern also holds in Turkish. Like Romeyka, the focussed constituent in Turkish is argued to be placed immediately preverbally, no matter what sub-type of focus it conveys (see Göksel and Kerslake, 2005; Kornfilt, 1997; Şener, 2010, i.a.). This is obligatory: nothing that bears information or contrastive focus can be placed in the postverbal field; it must occur immediately preverbally (see Erguvanli, 1984; Göksel and Kerslake, 2005; Kornfilt, 1997; Şener, 2010, i.a.).³ As a consequence, just like Romeyka, contrastive focus

cannot precede a contrastive topic in Turkish (since it no longer immediately precedes the verb) (see (72)):

(72) Turkish:

- a. Question:
Can'dan n'aber? O ne yedi partide?
'What about John? What did he eat at the party?'
Answers:
b. Valla Can-ı bil-mi-yor-um, ama ...
frankly Can-ACC know-NEG-PROG-1SG but
'Frankly, I don't know about John, but ...'
c. [Aylin]_{C-Top} [dolma-lar-dan]_{C-Foc} ye-di.
Aylin-NOM dolma-PL-ABL eat-PST-3SG
'Aylin ate from the dolmas.'
d. #[dolma-lar-dan]_{C-Foc} [Aylin]_{C-Top} ye-di.
dolma-PL-ABL Aylin-NOM eat-PST-3SG
'Aylin ate from the dolmas.'
(Şener, 2010: p. 19)

In the same way, as information focus must also be immediately preverbal, it must follow contrastive topics in Turkish too—again, parallel to Romeyka (see (73)):

(73) Turkish:

- a. Question:

between the focus and the verb (cf. Göksel and Özsoy, 2000; Güreş, 2020, among others). For example, *Ali* in (iB) and *I* in (iiB) below bear focus:

- (i) A: Yemek-ler-i kim pişir-di? (Ayşe mi Ali mi?)
dish-PL-ACC who cook-PST (Ayşe Q Ali Q)
'Who cooked the dishes? (Ali or Ayşe?)'
B: ALİ yemek-ler-i pişir-di, Ayşe değil.
Ali dish-PL-ACC cook-PST Ayşe not
'Ali cooked the dishes, not Ayşe.'
(ii) A: Kim ben-i çok sev-iyor?
who I-ACC much love-IPFV-1SG
'Who loves me so much?'
B: BEN sen-i çok sev-iyor-um.
I you-ACC much love-IPFV-1SG
'I love you so much.'

Indeed, this is possible, but, crucially, in this case the focalised constituent is the subject which appears higher than the object which occupies the position immediately to the left. So, when contrastive foci are subjects they cannot be adjacent to the verb because of the OV nature of the language whereby objects would have to occupy the immediate preverbal position (see Michelioudakis and Sitaridou (2020: p. 247) on the difficulty of diagnosing Multiple Wh-fronting in an OV language). Although we have seen in (67d) that this is not possible in Romeyka (and therefore, Romeyka and Turkish differ in this respect), similar effect is obtained in Romeyka when both the subject and object are new information foci (see Michelioudakis and Sitaridou, 2016: p. 9):
(iii) A: pios tinan endže?
who whom bring.PST.3SG
'Who brought (what) to whom?'

- (iv) B: o yusufis tin aişe yalemin endže.
the.NOM Yusufis the.acc Ayşe pencil bring.PST.3SG
'Yusuf brought a pencil to Ayşe.'

³ An anonymous reviewer noted that the immediate preverbal position is not the only position focus can appear in Turkish. It is possible to find non-preverbal focused constituents in Turkish with discourse-given elements intervening in

Filiz-in kardeş-ler-i ne iç-ti parti-de?
 Filiz-GEN sister-PL-POSS what drink-PST-3SG party-LOC
 'What did Filiz's sisters get to drink at the party?'

Answers:

Valla tüm kardeş-ler-den haberim yok, ama ...
 frankly all sister-PL-ABL news-POSS-1SG NEG but
 'Frankly, I do not know about all the sisters but ...'

b. [Filiz-in en küçük kardeş-i]_{C-Top} [rakı-dan]_{I-Foc} iç-ti.
 Filiz-GEN most young sister-3SG-
 POSS rakı-ABL drink-PST-3SG
 'Filiz's youngest sister drank (from the) rakı.'

c. #[rakı-dan]_{I-Foc} [Filiz-in en küçük kardeş-i]_{C-Top} iç-ti.
 rakı-ABL Filiz-GEN most young sister-3SG-
 POSS drink-PST-3SG
 'Filiz's youngest sister drank (from the) rakı.'
 (Şener, 2010: p. 35)

To summarise, Romeyka obeys the hierarchy in (74): topics always precede foci.

(74) Hierarchy of discourse features in Romeyka:

- a. Topic > Focus
- b. #Focus > Topic

Given our conclusion that there is a single focus position, and a single preverbal topic position (alongside a postverbal position for non-contrastive given information), this suggests the hierarchy of discourse features in (75):

(75) Articulation of discourse-related features in Romeyka:
 A-/C-TopicP I-/C-FocusP TP Given (non-contrastive)
 information

Given this hierarchy, [topic] and [focus] constitute autonomous structural notions in Romeyka: there is a dedicated left-peripheral projection encoding [topic], and another encoding [focus].

This represents a break from SMG, where the *ex situ*, preverbal position is generally associated with contrastive constituents (either topics and foci), but not information focus which favours a postverbal realisation. Instead, it appears that Romeyka has partially converged on the Turkish pattern: it exhibits the same immediately preverbal focus position regardless of focus type; and has the same restriction that foci must follow topics. However, when it comes to contrastive subject foci Romeyka deviates from Turkish either by having the contrastive focussed subject immediately to the left of the verb and the object higher or postverbally.

This pattern is highly suggestive, in that as it may reflect Romeyka's sociohistorical profile. Romeyka, as noted above, is the last Asia Minor Greek variety still spoken in Turkey; the speech community, by virtue of being Muslim, was exempted from the forced population exchange of 1923 which followed the cessation of the Greek-Turkish War (1919–1922). Consequently, it has undergone centuries-long contact with Turkish and concomitantly isolation from Standard Modern Greek. The parallels between the Turkish and Romeyka information structure systems may thus instantiate a contact effect.

The comparison with Pontic Greek is also informative in this regard. Pontic Greek is spoken primarily in Greece, and it too has been in some contact with Turkish prior to 1923 and significant contact (and thus attrition) with SMG since then. Interestingly, both Pontic varieties have converged in certain regards on the Turkish pattern, as noted: both have *ex situ* realisations for foci regardless of semantic type, for example. This may reflect parallel outcomes of Turkish contact. However, the organisation of their information structure also differs in significant ways: [contrast], for example, is an autonomous structural notion, with its own dedicated ContrastP projection realised by the *pa*-particle in Pontic Greek, whereas it seems to be a subfeature of [topic]/[focus] instead in Romeyka; and Pontic Greek does not observe the restriction that contrastive foci appear immediately preverbally. This may diagnose a difference in the contact profiles of Romeyka and Pontic Greek since their split (the latter to be taken the Islamisation onset, see Sitaridou, 2014a). The topic clearly awaits further investigation.

6 Conclusion

In this paper, we have sought to expand the coverage of formal work regarding the Asia Minor Greek variety, Romeyka, by investigating its word order and information structure. Regarding the latter, we have presented evidence that Romeyka patterns with its cognate variety, Pontic Greek, and Standard Modern Greek in having pragmatically unmarked, underlying VO word order, as well as verb-raising. However, Romeyka also exhibits frequent OV orders, attributable to information structural effects. As a result, we have argued that the organisation of Romeyka's information structure differs radically from SMG: topics and foci, of all semantic types, are realised *ex situ*, with no association with contrast (unlike SMG). We thus conclude that [topic] and [focus] are autonomous structural notions in Romeyka, realising heads in the clausal left periphery. Romeyka patterns instead with Turkish, which also has *ex situ* topics and foci, and like Romeyka limits them to the immediate preverbal position. Interestingly, the information structure in Romeyka's closest cognate variety, Pontic Greek, diverges in a significant way from both Romeyka and SMG in having a dedicated ContrastP projection (see Sitaridou and Kaltza, 2014) absent in Romeyka. The patterns in Romeyka and Pontic Greek may reflect subtly different patterns of contact with Turkish, though a more detailed investigation remains a goal for future inquiry (but see Sitaridou, 2022).

Data availability statement

The datasets presented in this article are not readily available because anonymisation of the data is not available yet. Requests to access the datasets should be directed to is269@cam.ac.uk.

Ethics statement

The study was approved by the Ethics Committee at the Faculty of Modern and Medieval Languages and Linguistics of the University of Cambridge ethics@mml.cam.ac.uk. The studies were conducted in accordance with the local legislation and institutional

requirements. Written informed consent was not required to participate in this study but all recordings were made ethically with the full oral consent of the speakers.

Author contributions

NN: Data Curation, Investigation, Writing – original draft. IS: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

References

- Agouraki, Y. (2010). It-clefts and stressed operators in the preverbal field of Cypriot Greek. *Lingua* 120, 527–554. doi: 10.1016/j.lingua.2008.10.011
- Alexiadou, A., and Anagnostopoulou, E. (1998). Parametrizing AGR: word order, V movement and EPP-checking. *Nat. Lang. Linguist. Theory* 16, 491–539. doi: 10.1023/A:1006090432389
- Alexiadou, A., and Anagnostopoulou, E. (2000). “Clitic Doubling and (Non-) Configurationality.” *Proceedings of NELS* 30, 17–28.
- Alexopoulou, T., and Baltazani, M. (2012). Focus in Greek wh-questions. In: Kučerová I, A Neeleman, eds. *Contrasts and Positions in Information Structure*. Cambridge University Press, 206–246.
- Beaver, D., and Clark, B. (2008). *Sense and sensitivity*. Oxford: Blackwell.
- Biberauer, T., and Roberts, I. (2010). “Subjects, Tense and verb-movement” in *Parametric variation: Null subjects in minimalist theory*. eds. T. Biberauer, A. Holmberg, M. Sheehan and I. Roberts (Cambridge: Cambridge University Press), 263–302.
- Büring, D. (2003). On D-trees, beans, and B-accents. *Linguist. Philos.* 26, 511–545. doi: 10.1023/A:1025887707652
- Büring, D. (2009). “Towards a typology of focus realization” in *Information structure*. eds. M. Zimmermann and C. Fery (Oxford: Oxford University Press), 177–205.
- Chatzikiyriakidis, S., Michelioudakis, D., and Spathas, G. (2015). *Greek focus operators and their associates. Studies for the Greek language*, vol. 35. Thessaloniki: Aristotle University of Thessaloniki, 167–179.
- Chomsky, N. (1972). *Studies on semantics in generative grammar*. The Hague: Mouton.
- Chomsky, N. (1982). *Some concepts and consequences of the theory of government and binding*. Cambridge, MA: MIT Press.
- Chomsky, N. (2001). “Derivation by phase” in *Ken Hale: a life in language*. ed. M. Kenstowicz (Cambridge, MA: MIT Press), 1–53.
- Dawkins, R. M. (1916). *Modern Greek in Asia minor: A study of the dialects of Silli, Cappadocia and Phárasa with grammar, texts, translations and glossary*. Cambridge: Cambridge University Press.
- Deffner, M. (1878). *Die Infinitive in den pontischen Dialekten und die zusammengesetzten Zeiten im Neugriechischen*. Berlin: Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin.
- Drettas, G. (1997). *Aspects pontiques*. Paris: Association de recherches pluridisciplinaires.
- Erguvanli, E. (1984). *The function of word order in Turkish grammar*. Berkeley: University of California Press.
- Frascarelli, M., and Hinterhölzl, R. (2007). “Types of topics in German and Italian” in *On information structure, meaning and form*. eds. K. Schwabe and S. Winkler (Amsterdam and Philadelphia: Benjamins), 87–116.
- Göksel, A., and Kerslake, C. (2005). *Turkish: A comprehensive grammar*. London: Routledge.
- Göksel, A., and Özsoy, A. S. (2000). “Is there a focus position in Turkish?” in *Studies on Turkish and Turkic languages*. eds. A. Göksel and C. Kerslake (Harrasowitz: Wiesbaden), 219–228.
- Gryllia, S. (2008). *On the nature of preverbal focus in Greek: a theoretical and experimental approach* (Vol. LOT dissertation series 200). Landelijke: Netherlands Graduate School of Linguistics.
- Gundel, J. K. (1998). On three kinds of focus. In P. Bosch and SandtR. van der, *Focus: linguistic, cognitive and computational perspectives* (pp. 293–305). Cambridge: Cambridge University Press.
- Gürer, A. (2020). *Information structure within interfaces: Consequences for the phrase structure*. Berlin: De Gruyter Mouton.
- Holmberg, A. (2005). Is there a little pro? Evidence from Finnish. *Linguist. Inq.* 36, 533–564. doi: 10.1162/002438905774464322
- Jackendoff, R. (1972). *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.
- Kisilier, M. (2007). *Word-order patterns in the Greek dialect of Mariupolis. Proceedings of the 3rd International Conference on Modern Greek Dialects and Linguistic Theory*, Nicosia.
- Koenenman, O., and Zeijlstra, H. (2014). The rich agreement hypothesis rehabilitated. *Linguist. Inq.* 45, 571–615. doi: 10.1162/LING_a_00167
- Kornfilt, J. (1997). On rightward movement in Turkish. In L. Johanson (Ed.), *He Mainz Meeting, Proceedings of the Seventh International Conference on Turkish Linguistics* (pp. 107–123). Wiesbaden: Harrasowitz.
- Krifka, M. (2007). “The semantics of questions and the focusation of answers” in *Topic and focus, crosslinguistic perspectives on meaning and intonation*. eds. C. Lee, M. Gordon and D. Büring (Dordrecht: Springer), 139–150.
- Lambrecht, K. (1994). *Information structure and sentence form*. Cambridge: Cambridge University Press.
- Mackridge, P. (1987). *Greek-speaking Moslems of north-East Turkey: prolegomena to a study of the ophitic sub-dialect of Pontic*. *Byz. Mod. Greek Stud.* 11, 115–137. doi: 10.1179/030701387790203037
- Mackridge, P. (1999). “The Greek spoken in the region of Pontus” in *Dialect enclaves of the Greek language*. eds. A. F. Christidis, M. Arapopoulou and G. Giannouloupoulou (Athens: Hellenic Republic, Ministry of National Education and Religious Affairs), 101–105.
- Melanofrydis, P. (2001). *Οι Κλωστοί. Θεσσαλονίκη: Αδελφοί Κυριακίδη*.
- Michelioudakis, D., and Sitaridou, I. (2012). “Syntactic microvariation: dative constructions in Greek” in *Datives in variation: A micro-comparative perspective*. eds. R. Etxepare and B. Fernández (Oxford: Oxford University Press), 212–255.
- Michelioudakis, D., and Sitaridou, I. (2013). Multiple wh-fronting in Romyekia. In A. Ralli, B. Joseph and M. Janse (Ed.), *Proceedings of the Fifth International Conference of Modern Greek Dialects and Linguistic Theory (MGDLT5)*, Ghent, 20–22 September 2012, (pp. 353–378).
- Michelioudakis, D., and Sitaridou, I. (2016). Recasting the typology of multiple wh-fronting: evidence from Pontic Greek. *Glossa* 1, 40–72. doi: 10.5334/gjgl.72
- Michelioudakis, D., and Sitaridou, I. (2020). “Towards a formal model of transfer under contact: contrasting Asia minor Greek to mainland Greek and Turkish in search of syntactic borrowings” in *Theoretical approaches to contrastive linguistics: Morphological and syntactic perspectives*. eds. M. Georgiamentis, G. Giannouloupoulou, M. Koliopoulou and A. Tsokoglou (London: Bloomsbury Studies in Theoretical Linguistics), 246–261.

Conflict of interest

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.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Molnár, V. (2006). "On different kinds of contrast" in *Architecture of focus. Studies in generative grammar*, eds. V. Molnár and S. Winkler, vol. 82 (Berlin/New York: Mouton de Gruyter), 197–233.
- Neeleman, A., and van de Koot, H. (2008). Dutch scrambling and the nature of discourse templates. *J. Comp. Ger. Linguist.* 11, 137–189. doi: 10.1007/s10828-008-9018-0
- Neocleous, N. (2020). Word order and information structure in Romeyka: A syntax and semantics interface account of order in a minimalist system: PhD dissertation, University of Cambridge.
- Neocleous, N. (2022). "The evolution of VO and OV alternation in Romeyka" in *Word order variation – Semitic, Turkic, Caucasian and indo-European languages in contact*. STUF – language typology and universals. In Jügel, Thomas, and Hiwa Asadpour, eds., *Word Order Variation: Semitic, Turkic and Indo-European Languages in Contact*, De Gruyter, Berlin, Germany.
- Neocleous, N., and Sitaridou, I. (2022). Never just contact. The rise of final auxiliaries in Asia minor Greek. *Diachronica* 39, 369–408. doi: 10.1075/dia.17048.neo
- Özkan, H. (2013). The Pontic Greek spoken by Muslims in the villages of Beşköy in the province of present-day Trabzon. *Byz. Mod. Greek Stud.* 37, 130–150. doi: 10.1179/0307013112Z.000000000023
- Parcharidis, I. (1880). Γραμματική τῆς διαλέκτου Τραπεζοῦντος: Περὶ ἐπιθέτων, ἀντωνομιῶν καὶ βαροτόνων ρημάτων, vol. *Χειρόγραφο* 335. Athens: Κέντρο Ἑρεῦνης τῶν Νεοελληνικῶν Διαλέκτων καὶ Ἰδιομάτων, Ἀκαδημία Ἀθηνῶν.
- Pollock, J.-Y. (1989). Verb-movement, universal grammar, and the structure of IP. *Linguist. Inq.* 20, 365–424. Stable URL: <http://www.jstor.org/stable/4178634>
- Roberts, I. (1993). *Verbs and diachronic syntax: A comparative history of English and French*. Dordrecht: Kluwer.
- Rooth, M. E. (1992). A theory of focus interpretation. *Nat. Lang. Semant.* 1, 75–116. doi: 10.1007/BF02342617
- Sağlam, E. (2017). Constitutive ambiguities. Subjectivities and memory in the case of Romeyka-speaking communities of Trabzon: Unpublished PhD dissertation, Birkbeck, University of London.
- Schreiber, L., and Sitaridou, I. (2017). Assessing the sociolinguistic vitality of Istanbulite Romeyka: an attitudinal study. *J. Multiling. Multicult. Dev.* 39, 1–16. doi: 10.1080/01434632.2017.1301944
- Şener, S. (2010). (Non-)peripheral matters in Turkish syntax: Unpublished PhD dissertation, University of Connecticut.
- Sitaridou, I. (2013). "Greek-speaking enclaves in Pontus today: the documentation and revitalization of Romeyka" in *Keeping languages alive*. 322 word order and information structure in Romeyka: A syntax and semantics interface account of order in a minimalist system language endangerment: Documentation, pedagogy and revitalization. eds. M. C. Jones and S. Ogilvie (Cambridge: Cambridge University Press), 98–112.
- Sitaridou, I. (2014a). The Romeyka infinitive: continuity, contact and change in the Hellenic varieties of Pontus. *Diachronica* 31, 23–73. doi: 10.1075/dia.31.1.02sit
- Sitaridou, I. (2014b). Modality, antiverdicality and complementation: the Romeyka infinitive as a negative polarity item. *Lingua* 148, 118–146. doi: 10.1016/j.lingua.2014.05.017
- Sitaridou, I. (2015). Reframing the Phylogeny of Asia Minor Greek: The View from Pontic Greek" CHS Research Bulletin 4, no.1. http://nrs.harvard.edu/urn-3:hnc:essay:SitaridouI.Reframing_the_Phylogeny_of_Asia_Minor_Greek.2016.
- Sitaridou, I. (2017a). "The rise of novel negators in Romeyka: negative existential cycle and monoclausality", Workshop on the negative existential cycle from a historical-comparative perspective, Stockholm University, 04–05/05/2017.
- Sitaridou, I. (2017b). "Null objects in Romeyka: contact with Turkish or internal conditioning?", Talk at the Workshop on historical language contact in English and beyond, Aristotle University of Thessaloniki, 31/03–02/04/2017.
- Sitaridou, I. (2021). "(in)vulnerable inflected infinitives as complements to volitionals and modals: evidence from Galician and Romeyka" in *Syntactic features and the limits of syntactic change*. eds. T. Eythrrsson and J. G. J. Nsson (Oxford: Oxford University Press), 161–177.
- Sitaridou, I. (2022). "On the redundancy of a theory of language contact: cue-based reconstruction in a sociolinguistically informed manner" in *Language change: Theory and methodologies in the 21st century*. eds. N. Lavidas and K. Nikiforidou (Leiden: Brill), 15–52.
- Sitaridou, I. (2023a). "Romeyka" in *Endangered Languages in Turkey*. eds. M. Bagriacik, O. Demirok and B. Ozturk (Istanbul: The Laz Institute), 118–146.
- Sitaridou, I. (2023b). Romeyka. "Türkiye'nin tehlike altındaki dilleri" içinde, Metin Bağrıaçık and Ömer Demirok ve Balkız Öztürk (Ed.), İstanbul: Laz Enstitüsü.
- Sitaridou, I. (2023c). "Key Insights about Romeyka's Linguistic Description and Evolution, Vitality, and Revitalisation Prospects After 15 Years of Field-Based Research". *Translated by Jerneja Kavčič. Keria: Studia Latina Et Graeca* 25, 9–32. doi: 10.4312/keria.25.2.9-32
- Sitaridou, I. (2024). Τα ρομέικα χθες και σήμερα. Romeyka exhibition catalogue. Octana: Thessaloniki, 26–54.
- Sitaridou, I. (in prep.). The evolution of negation in Romeyka: Ms, University of Cambridge.
- Sitaridou, I., and Kaltsa, M. (2014). Contrastivity in Pontic Greek. *Lingua* 146, 1–27. doi: 10.1016/j.lingua.2014.04.005
- Tsimpli, I. M. (1995). "Focusing in modern Greek" in *Discourse configurational languages*. ed. K. E. Kiss (Oxford: Oxford University Press), 176–207.
- van der Wal, J. (2016). Diagnosing focus. *Stud. Lang.* 40, 259–301. doi: 10.1075/sl.40.2.01van
- Vikner, S. (1995). *Verb movement and expletive subjects in the Germanic languages*. Oxford: Oxford University Press.
- Vikner, S. (1997). "V-to-I movement and inflection for person in all tenses" in *The new comparative syntax*. ed. L. Haegeman (London: Longman), 189–213.

Glossary

ABL - ablative

ACC - accusative

FUT - future

GEN - genitive

IMP - imperative

IPFV - imperfective

LOC - locative

MOD - modal

NEG - negation

NOM - nominative

PRT - particle

PL - plural

POSS - possessive

PST - past

SG - singular

A-TOP - aboutness topic

C-Topic - contrastive topic

C-FOC - contrastive focus

DP - determiner phrase

FOC - focus

I-FOC - information focus

NP - noun phrase

PG - Pontic Greek

O - object

SMG - Standard Modern Greek

S - subject

TOP - topic

TP - tense phrase

V - verb

VP - verb phrase

vP - light verb phrase

Frontiers in Psychology

Paving the way for a greater understanding of human behavior

The most cited journal in its field, exploring psychological sciences - from clinical research to cognitive science, from imaging studies to human factors, and from animal cognition to social psychology.

Discover the latest Research Topics

[See more →](#)

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact

