Formal approaches to complexity in heritage languages

Edited by

Maria Polinsky

Michael T. Putnam

Current Issues in Bilingualism

Current Issues in Bilingualism

In this series:

1. Shim, Ji Young. OV and VO variation in code-switching.
2. Rodríguez González, Eva & Rosita L. Rivera (eds.). Language assessment in multilingual settings: Innovative practices across formal and informal environments.

ISSN (print): 2747-9919
ISSN (electronic): 2747-9927

Formal approaches to complexity in heritage languages

Edited by

Maria Polinsky

Michael T. Putnam

Maria Polinsky & Michael T. Putnam (eds.). 2024. *Formal approaches to complexity in heritage languages* (Current Issues in Bilingualism). Berlin: Language Science Press.

This title can be downloaded at:
redefine\lsURL

ISBN: no digital ISBN
no print ISBNs!
ISSN (print): 2747-9919
ISSN (electronic): 2747-9927
no DOI

Cover and concept of design: Ulrike Harbort
Fonts: Libertinus, Arimo, DejaVu Sans Mono
Typesetting software: XeLaTeX

redefine \publisherstreetaddress
redefine \publisherurl

Storage and cataloguing done by redefine \storageinstitution

Contents

Chapter 1

Linguistic complexity in heritage languages: An introduction

Maria Polinsky[a], Michael T. Putnam[b,c] & Joe Salmons[d]
[a]University of Maryland, College Park [b]Penn State University [c]University of Greenwich (CREL) [d]University of Wisconsin – Madison

"I know it when I see it." Justice Potter Stewart's famous 1964 sentence referred to a much more exciting phenomenon than the one addressed in this volume (Gewirtz 1996), but it is probably the best distillation of the way complexity in language has been approached. Complexity seems easy to observe, but it lacks clearly established parameters, and its definition remains incredibly subjective and variable. Things only get more complicated (no pun intended) when this notion, however undefined, is used by different researchers to explain sets of language phenomena they observe.[1] In particular, complexity often emerges as a prominent explanation in work on language acquisition, bilingualism, and language change. What all these research domains have in common is that they consider and compare two varieties of what is arguably the same language and strive to explain how these varieties differ: adult/child language in L1 acquisition; native-like/non-native like versions in second language acquisition; monolingual/bilingual varieties of one and the same language in bilingualism, and finally, two temporal slices of one and the same language in language diachrony. The comparisons generally rely on complexity as an explanatory tool, but that tool itself takes very different shapes in different hands. Indeed, valiant attempts

[1]See Shannon (1948a,b) on information theory, and consider the inclusion of measures of entropy and surprisal in linguistic research (e.g., Levy 2008). We will not comment further on these more general treatments of complexity as defined and measured in terms of information theory.

Maria Polinsky, Michael T. Putnam & Joe Salmons. 2024. Linguistic complexity in heritage languages: An introduction. In Maria Polinsky & Michael T. Putnam (eds.), *Formal approaches to complexity in heritage languages*, 1–13. Berlin: Language Science Press. DOI: ??

to define complexity have been made by a number of researchers (consider in particular Culicover 2013, Dahl 2004, Miestamo et al. 2008, Trudgill 2011).

We will not be able to resolve these debates here, but this volume seeks a common core in the existing treatments of complexity, and we apply this common core to heritage languages and their speakers, an area where complexity is constantly discussed, if not always in satisfying ways. That is, we focus on issues of bilingualism, and through it, language acquisition and language change. All the instances of bilingualism considered in this volume fall under the rubric of *unbalanced bilingualism*, where the bilingual's home language becomes subordinate to the societal language in adulthood. The former are widely referred to as *heritage languages*, and its speakers, as *heritage speakers*. A common view of a heritage language relies on two main criteria: it is acquired as an L1 and it is not the dominant language of the larger (national) society (Rothman 2009, Montrul 2016, Polinsky 2018, Wintai et al. 2021, and references therein). The social context of heritage languages as minority languages is necessary to understanding the limited social domains where heritage languages may be acquired, used, and maintained. While this volume is not concerned with the social issues of heritage language use, it is important to underscore that the social context is conducive to a variety of terms used to describe such languages as "minority", "minoritized", or "non-dominant".

Focusing on structural properties of heritage languages, several comparisons can be made within a dominant/heritage bilingual dyad (see Scontras & Putnam 2020 on the importance of the dyads in heritage language research):[2]

(i) between the dominant language as spoken by a monolingual and by a heritage speaker (e.g., monolingual Canadian English and Canadian English spoken by a heritage speaker of Korean living in Toronto);

(ii) between a heritage language and the corresponding homeland language (e.g., heritage Korean as spoken in Toronto and Korean spoken in Seoul);

(iii) between a heritage language and the baseline: the language of input, which is typically the language spoken by the first-generation adult immigrants (e.g., Korean as spoken by 2+ generations of descendants of Korean immigrants and Korean as spoken by the first generation of Korean immigrants all residing in Toronto)

[2]In our illustrative examples, we rely on the ongoing work of the HLVC project: Heritage Language Variation and Change in Toronto, https://ngn.artsci.utoronto.ca/HLVC/0_0_home.php

An evaluation of dominant-language varieties (i) lies beyond the scope of this volume, and we will not comment on it here. Comparisons (ii) and (iii) are not always equally possible. Sometimes data on the baseline are simply missing, which makes an investigation along the lines of (iii) impossible. Sometimes there is no homeland, which makes comparisons in (ii) tricky. Consider here the example of many indigenous minority languages whose traditional language variety is no longer spoken (Grenoble & Osipov 2023, Kehayov 2017, Polinsky 2018: Ch. 1, Sasse 1992, among others), often resulting in populations of receptive bilinguals (Holmes & Putnam 2020, Sherkina-Lieber 2015), or the situation of a dialect that originated in the homeland but is no longer spoken there, such as Pomeranian, an East Low German language now moribund in Europe and North America but spoken actively in Brazil with growing state support (Savedra 2020).

In early heritage language studies, most comparisons were made between a given heritage language and a homeland language ((ii) above), but as the understanding of heritage bilingualism progressed and the empirical base of heritage-language studies grew, comparisons with the baseline or three-way comparisons (homeland/baseline/heritage variety) have grown more common. Most papers in this volume pursue such a line of inquiry, and this is one area where the notion of complexity becomes prominent.

Some folk perceptions of complexity are easy to dispense with, for example, the idea that some languages are more complex than others – one habitually hears something like "Hungarian is more complex than English" or "Italian is less complex than German". These are often reflections of subjective difficulties in learning a second language in adulthood, through structured instruction, and their echoes in research work on complexity are not too loud. Equally obvious is the idea that complexity has to be defined on multiple levels: phonetics and abstract phonological representations, morphology and morphosyntax, semantics, or the lexicon. The number of dialects and registers in a given language may also contribute to complexity.

While these misconceptions are easy to set aside, another folk view is prominent in work on bilingualism and this one is more insidious: the view of heritage languages as simplified as compared to their traditional baselines. One of the most prevalent conceptions about heritage languages is that they are somehow "simpler" and "less complex" compared with monolingual grammars. Such a view of heritage languages and their speakers can, and does, have detrimental effects on the public perception of heritage bilingualism, pedagogical initiatives and curriculum development, as well as the development and advancement of theoretical and experimental studies of these grammars. This view faces critical scrutiny in the chapters in this volume, which show that changing complexity in

heritage languages is often a trade-off between different levels of representation (see especially Laleko and Varatharaj et al., both this volume).

To advance our understanding of complexity and how it can enhance linguistic analysis, we distinguish between the defining properties of complexity as a phenomenon, diagnostics of complexity, and the ways complexity is modeled, measured or operationalized in language sciences. In what follows, we will discuss these facets of complexity in turn.

In defining complexity as a phenomenon, researchers distinguish between language-specific complexity and global complexity, best viewed through the lens of information theory (the latter approach is prominent in Varatharaj et al., this volume). Regardless of whether complexity is defined internally to language or globally, one can further distinguish between complexity in the signal (form) and in the content of the signal (meaning). On the signal plane, it is common to consider those entities more complex that include more constituent parts, but the moment such an approach is assumed, implicitly or explicitly, the question arises as to what counts as a unit and a constituent. To make things somewhat more precise, many approaches to complexity distinguishes between abstract (grammatical) atoms, rules that relate such atoms to units of form, and rules that relate such atoms to units of meaning. Distributed Morphology (see Harley & Noyer 1999, McGinnis-Archibald 2016 for overviews) and Nanosyntax (Caha 2009) are good examples of models that address these aspects of complexity as a purely linguistic phenomenon, and in this volume, Lohndal and Putnam is an example of such an approach (see also Lohndal & Putnam 2021).

Assuming a division of labor between abstract units, rules, exponents, and interface conditions, the main finding concerning bilingual systems has to do with a reduction or contraction of those domains where interface conditions are applied, where prime examples have to do with the interface between syntax and discourse (see Laleko 2021 for an overview) and the reduction or regularization in the actual inventory of exponents (e.g. in the constantly discussed reduction of case and other morphological marking). If this approach to complexity is on the right track, then it behooves us to distinguish between units of computation, rules of computation, and interfaces. As is commonly noted, changes occur primarily at the interfaces and in the reduction of computational domains, whereas the abstract units and the rules operating on them remain stable regardless of language contact.

A related approach in some sense is Culicover's proposal for two kinds of complexity. The first, *formal complexity*, is "the measure of the amount of idiosyncrasy in a grammar" and the second, *processing complexity*, is "a measure of the resources required to compute the correspondences between particular

grammatical forms and their meanings" (2013:3). Current theorizing no longer treats these two forms of complexity in isolation, allowing us to envisage a path forward where learning to parse (roughly) equates to acquiring and fine-tuning abstract structure. Although the question of which form of complexity comes first and guides the other remains a matter of intense debate (cf. Christiansen & Chater 2016, Lightfoot 2020), scholars on both sides see the benefit of involving both *formal* and *processing complexity* in linguistic analysis.

Decomposing complexity and looking for its properties at the local or global level remains a rather uncommon strategy. A more common approach to complexity is to look for its telltale signs, and in that regard, one finds parallels between research on complexity in bilingual systems and in child language acquisition. Common measures of complexity include frequency differences, with less frequent items or phenomena associated with greater complexity. For example, it is often assumed that passive-voice constructions are more complex than their active counterparts, or that closed syllables are more complex than CV structures. But connecting frequency and complexity creates a circle of a kind: are structures more complex because they are less frequent, or are they less frequent because they are complex? Since the causal relation is unclear, one can use frequency as a possible diagnostic of complexity but not as an explanation of what can be complex.

Cross-linguistic distribution is another way of diagnosing complexity, the idea being that more complex structures have more restricted distributions. If something is cross-linguistically common, such commonality is often associated with naturalness – another vague but widely used notion. Both cross-linguistic frequency and naturalness are often used in studies of sound systems, where natural sound patterns are grounded in physical properties of speech, while unnatural sound patterns arguably have no such physical basis (cf. Blevins 2004, 2007). The use of cross-linguistic distribution as a metric of complexity and naturalness in sound systems has been valuable, but it is less clear how to apply similar criteria beyond phonetics; once the connection to physical properties of speech is removed, the evaluation metric becomes less clear.

Naturalness and cross-linguistic distribution are also associated with ease of acquisition, and with that a connection between complexity and L1 learning. The age or order of acquisition is also associated with greater complexity; for example, in the L1 acquisition of tone in Mandarin, T3 (also known as the dip(ping) tone) is acquired later than T1, T2, or T4 (Li & Thompson 1977). Likewise, this tone is particularly challenging to L2 learners and heritage speakers of Mandarin (Chan & Chang 2019). Unlike the other tones, this one can be decomposed into fall and subsequent rise, thus being arguably more complex than the other tones.

It is a common assumption that the three signs of complexity discussed here align: less frequent structures take longer to acquire and may be less common cross-linguistically. Things are actually more nuanced, and we would like to offer several considerations. First, the correlations between frequency, age of acquisition, and cross-linguistic distribution are not iron-clad. Some phenomena that are cross-linguistically rare are unproblematic in acquisition. Subject-verb inversion in English questions is an example (see Grinstead et al. 2018); such inversion is typologically uncommon, but its acquisition goes without a hitch because the inversion is tied to the fundamental notion of finiteness. Second, the connection between frequency and acquisition is predicated on the idea that acquisition is based on lots of data which then generalize on the basis of similarity. On such a view, abstract knowledge comes after a process of data collection and statistical compression. But there is also the alternative view according to which abstract knowledge guides the use of input and shapes the form of grammatical knowledge throughout development (e.g., Lidz 2018). In this view, the patterns of data that children use to acquire grammar differ from the actual structures responsible for those patterns, and in acquiring a language, statistical information in the environment is used for inference but not for building the representations. Next, the cross-linguistic distribution of patterns is the weakest of this set of complexity diagnostics, as our understanding of what is and is not common across languages depends on the extent of our empirical knowledge (after all there are over seven thousand languages and we know about just a small fraction of that sample) as well as social, political or demographic factors of language distribution.

A related and perhaps equally difficult notion is "markedness", with appeal to similar criteria and tests, including that unmarked features are more "prototypical" (see Haspelmath 2006 on the multiple, at times contradictory, understandings of this notion). Battistella (1990: 1) defines *marked* and *unmarked* this way: in terms of polar oppositions in language "the simpler, more general pole is the unmarked term of the opposition while the more complex and focused pole is the marked term". For instance, Russian /b,d,g/, etc. are specified for or marked as "voiced", [voice], while /p,t,k/, etc. are unspecified on privative views of phonological features or otherwise "voiceless" [–voice]. The presence of such a feature is supported by the spread of voicing in clusters, so that *mók* surfaces as [mógbʲi] 'were s/he getting wet' (Dresher 2009: 93, example ultimately from Halle 1971: 22), as well as its deletion in codas, a prosodically weak position. Many other oppositions do not show obviously simple vs. complex forms, e.g. semantic pairs like *boy/girl*. Battistella is not particularly bothered that "[n]o single diagnostic is a fully reliable indicator" of markedness for any given pair, concluding that "The fact that we cannot define the notions marked and unmarked perfectly is no more

surprising than the fact that we cannot define the notion verb perfectly" (Battistella 1990: 45). In their contribution to this volume, D'Alessandro and Terenghi address the connection between markedness and complexity, and propose an important distinction between the two: while markedness is typically assessed with respect to an individual item/language segment, complexity is a more global concept, one that covers a collection of items or language segments. Removing markedness from one of the items in a paradigm may lead to increased complexity elsewhere, hence the recurrent trade-off in complexity across language levels. D'Alessandro and Terenghi's own proposal, couched in terms of Linguistic Minimalism, relies on the recognition of grammatical features, defined as properties necessary and sufficient to identify grammatical objects or categories (cf. Adger & Svenonius 2012 on features in Minimalist syntax).

Let us turn now to the content of the individual contributions, less for a traditional summary of contents than to note what defining properties they deal with, language specific or global, what diagnostics or model they use. These chapters are different and approach complexity in unique and varied ways, for instance in employing various experimental approaches, ranging from elicited productions to judgment tasks (Kpogo et al., Rinke et al., Laleko), and distinct kinds of cross-linguistic comparisons (Lohndal & Putnam, Rinke et al., Varatharaj et al., D'Alessandro & Terenghi).

With an eye on diagnostics of complexity, Felix Kpogo, Alexandra Elizabeth Kohut, & Charles Chang show that second generation Twi speakers in the U.S. prefer a syntactic over a morphological strategy for diminutive formation, in "Expressing diminutive meaning in Twi: The role of complexity and language-specific preferences". They argue that the morphological strategy is "incrementally" more complex and that the younger generation minimizes that complexity. They see this as a local and not a universal strategy, since it does not appear to match broader diachronic or typological patterns.

Esther Rinke, Cristina Flores, & Jacopo Torregrossa's "How different types of complexity can account for difficult structures in bilingual and monolingual language acquisition" also emphasizes diagnostic issues, with attention to age of acquisition. They investigated 180 European heritage speaking-children in Switzerland, which were divided into three different and distinct environmental languages (e.g., French, German, and Italian). The main contribution of Rinke et al.'s chapter lies in showing that the difficulties encountered by these bilingual children cannot be accounted for in terms of a single notion of complexity. Rather, individual structures and phenomena may be complex in acquisition for different reasons. In particular, Rinke et al. identified four main notions of

complexity that are related to the structures they investigated: derivational complexity; memory-based learning; context dependency of rules, and multiple form-function mappings.

In "The complexity of word order change in a flexible system: On stability and variation in heritage Russian word order", Oksana Laleko points to the need for nuanced diagnostics in assessing complexity. English-dominant heritage speakers of Russian do not show any general move toward English-like SVO structures, nor difficulties with information structure. Instead, they provide evidence for stronger SOV patterns, distinctly unlike English. She characterizes the use of SOV patterns as "a rarely documented case of non-transfer-induced complexity-preserving change".

Terje Lohndal & Michael T. Putnam's contribution, "Expanding structures while reducing mappings: Morphosyntactic complexity in agglutinating heritage languages", emphasizes the importance of agglutinating and polysynthetic languages for more general claims concerning modeling complexity in heritage languages. Building on their previous work (Lohndal & Putnam 2021) and a related architecture developed by López (2020), Lohndal & Putnam argue for modeling morphological complexity in heritage languages through the lens of a late-insertion, exoskeletal model of the syntax-morphology interface. A point of emphasis is the tension between lexicalizing larger spans of syntactic structure, while also "shrinking" computational domains (see also Scontras et al. 2018 and Polinsky & Scontras 2020 for similar arguments).

Ashvini Varatharaj, Gregory Scontras, & Naomi Nagy, in "A multi-generational analysis of heritage language complexity", emphasize global defining properties across generations of heritage speakers of a half dozen heritage languages in Toronto. They explore possible "complexity trade-offs" between morphological and syntactic complexity over generations. They find some support for a shift from morphological to word-order complexity. This is especially so for Ukrainian speakers, perhaps, they argue, due to the very different morphological systems of English and Ukrainian.

In "Non-monotonic functional sequences: A new metric for complexity in heritage languages", Roberta D'Alessandro & Silvia Terenghi focus on diagnostics, seeking motivation in language design and thus also framing complexity as a global phenomenon. As hinted at above, complexity in part needs to be reckoned not in terms of exponents but in terms of 'monotonicity and uniformity' in computation. Once monotonicity and uniformity are lacking, they argue, that opens the door to language change.

There are clearly common threads running through these papers. For example, Laleko, Kpogo et al., Rinke et al., and Varatharaj et al. all understand complex-

ity as something that has to be considered in terms of a specific level of representation or even construction. Of course, the papers arrive at no consensus beyond the already-familiar (if hardly universally accepted!) understanding that language contact, language change, and bilingualism – here specifically under the circumstances of heritage language bilingualism – do not lead inevitably to simplification, or even to complexification, but rather can involve "complexity invariance", to follow Laleko (this volume).

We summarize the proposed set of defining features of complexity in the Figure 1 below. We use the traditional linguistic contrast between form and meaning to capture facets of complexity on the language-specific level, and its more general counterpart, the contrast between signal and content on the global level.

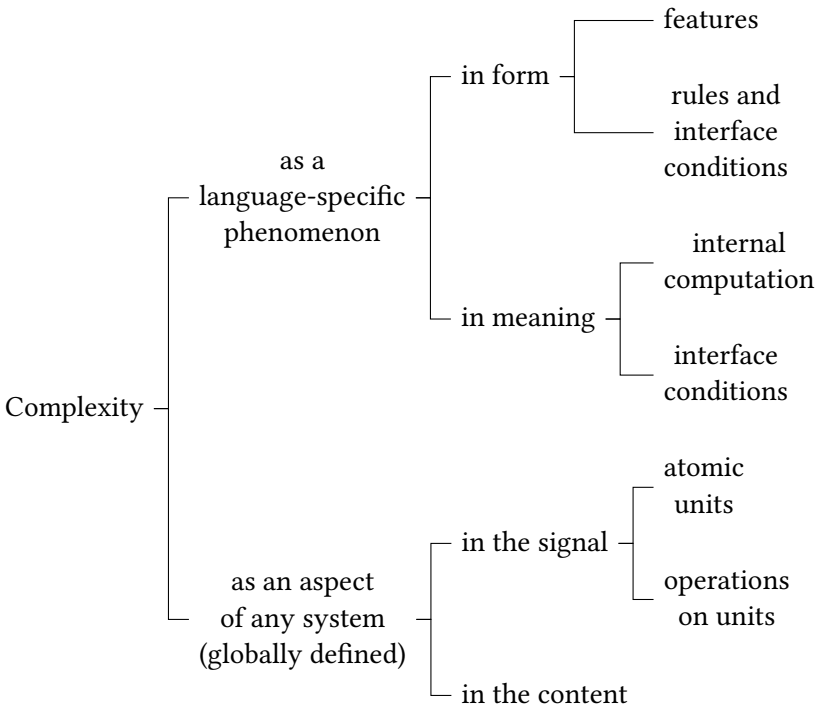


Figure 1: Defining features of complexity

Building on the representation in Figure 1, while most contributions in this volume define complexity as a global, domain-general phenomenon, Kpogo et al. propose that certain instances of complexity, such as the syntactic (vs. the morphological) formation of diminutives in Twi by second generation heritage

speakers, can be viewed as something language-specific under certain conditions. Complexity is clearly present in ambient input according to the studies set forth by Rinke et al. and Varatharaj et al., while Kpogo et al., D'Alessandro & Terenghi, Laleko, and Lohndal & Putnam find complexity in the content of representations. The findings of Varatharaj et al. once again emphasize the importance of typological proximity and distance between the languages existing within a dyad (Putnam et al. 2018, Scontras & Putnam 2020). Aside from Kpogo et al., most contributions investigate measures of complexity as they relate to forms (i.e., representations) rather than meaning. D'Alessandro & Terenghi and Lohndal & Putnam conceptualize the notion of complexity in abstract feature rules while also holding to some notion of modularity and interface compatibility. Kpogo et al., Laleko, and Rinke et al.'s contributions emphasize the importance of language-internal restructuring, or internal computation. Varatharaj et al.'s proposal of 'complexity trade-offs' straddles this distinction. Laleko's study combines the requirement of interface conditions and internal computation, resulting in word-order patterns that are distinctly complexity-preserving.

In closing, the diversity of definitions of complexity and their implementations in analysis found in this volume raise more questions than they ultimately resolve. Nonetheless, we interpret the proposals and supporting arguments for various strands of complexity as an exciting opportunity to further engage with this central notion in heritage language bilingualism. We eagerly look forward to seeing where these discussions lead and how they may direct and shape additional studies moving forward.

Acknowledgments

The editors would like to thank the authors of these chapters and their reviewers for their work. We would also like to thank the editors of the Language Science Press series *Current Issues in Bilingualism* for making this volume happen. We are grateful to Terje Lohndal, Naomi Nagy, Esther Rinke, and Greg Scontras for comments on the earlier version of this introduction. All errors are our responsibility.

References

Adger, David & Peter Svenonius. 2012. Features in minimalist syntax. In Cedric Boeckx (ed.), *The Oxford handbook of linguistic minimalism*, 27–51. Oxford: Oxford University Press.

Battistella, Edwin. 1990. *Markedness: The evaluative superstructure of language*. Albany: State University of New York Press.

Blevins, Juliette. 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge, UK: Cambridge University Press.

Blevins, Juliette. 2007. The importance of typology in explaining recurrent sound patterns. *Linguistic Typology* 11. 107–113.

Caha, Pavel. 2009. *The nanosyntax of case*. Tromsø: University of Tromsø. (Doctoral dissertation). https://hdl.handle.net/10037/2203.

Chan, Lei & Charles Chang. 2019. Perception of nonnative tonal contrasts by Mandarin-English and English-Mandarin sequential bilinguals. *Journal of the Acoustical Society of America* 146(2). 956–972. DOI: 10.1121/21.512052.

Christiansen, Morten H. & Nick Chater. 2016. *Creating language: Integrating evolution, acquisition, and processing*. Cambridge, MA: MIT Press.

Culicover, Peter W. 2013. *Grammar and complexity: Language at the intersection of competence and performance*. Oxford: Oxford University Press.

Dahl, Östen. 2004. *The growth and maintenance of linguistic complexity*. Amsterdam: John Benjamins.

Dresher, B. Elan. 2009. *The contrastive hierarchy in phonology*. Cambridge: Cambridge University Press.

Gewirtz, Paul. 1996. On 'I Know It When I See It.' *The Yale Law Journal* 105. 1023–1047.

Grenoble, Lenore A & Boris Osipov. 2023. Understanding language shift. *Linguistic Approaches to Bilingualism* 13(1). 122–132. DOI: 10.1075/lab.22100.gre.

Grinstead, John, Maria Vega-Mendoza & Grant Goodall. 2018. Inversion and finiteness in Spanish and English: Developmental evidence from the optional infinitive and optional inversion stages. *Language* 94. 575–610.

Halle, Morris. 1971. *The sound pattern of Russian: A linguistic and acoustical investigation*. The Hague: Mouton & Co.

Harley, Heidi & Ralph Noyer. 1999. State-of-the-article: Distributed Morphology. *GLOT International* 4. 3–9.

Haspelmath, Martin. 2006. Against markedness (and what to replace it with). *Journal of Linguistics* 42(1). 25–70. DOI: 10.1017/S0022226705003683.

Holmes, Bonnie C. & Michael T. Putnam. 2020. Mental representations in receptive multilingual grammars. *Linguistic Approaches to Bilingualism* 10(3). 309–314.

Kehayov, Petar. 2017. *The fate of mood and modality in language death: Evidence from minor finnic*. Berlin: Mouton de Gruyter.

Laleko, Oksana. 2021. Discourse and information structure in heritage languages. In Silvina Montrul & Maria Polinsky (eds.), *The Cambridge handbook of heritage languages and linguistics*, 691–727. Cambridge, UK: Cambridge University Press.

Levy, Roger. 2008. Expectation-based syntactic comprehension. *Cognition* 3(106). 1126–1177.

Li, Charles N. & Sandra Thompson. 1977. The acquisition of tone in Mandarin-speaking children. *Journal of Child Language* 4. 185–199.

Lidz, Jeff. 2018. The scope of children's scope: representation, parsing and learning. *Glossa* (1). 33. DOI: 10.5334/gjgl.339.

Lightfoot, David W. 2020. *Born to parse: How children select their languages*. Cambridge, MA: MIT Press.

Lohndal, Terje & Michael T. Putnam. 2021. The tale of two lexicons: Decomposing complexity across a distributed lexicon. *Heritage Language Journal* 18(2). 1–29.

López, Luis. 2020. *Bilingual grammar*. Cambridge, UK: Cambridge University Press.

McGinnis-Archibald, Martha. 2016. Distributed morphology. In Andrew Hippisley & Gregory Stump (eds.), *The cambridge handbook of morphology*, 390–423. Cambridge: Cambridge University Press.

Miestamo, Matti, Kaius Sinnemäki & Fred Karlsson (eds.). 2008. *Language complexity: Typology, contact, change*. Amsterdam: John Benjamins.

Montrul, Silvina. 2016. *The acquisition of heritage languages*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139030502.

Polinsky, Maria. 2018. *Heritage languages and their speakers*. Cambridge, UK: Cambridge University Press. DOI: 10.1017/9781107252349.

Polinsky, Maria & Gregory Scontras. 2020. Understanding heritage languages. *Bilingualism: Language and Cognition* 23(1). 4–20. DOI: 10 . 1017 / S1366728919000245.

Putnam, Michael, Matt Carlson & David Reitter. 2018. Integrated, not isolated: Defining typological proximity in an integrated multilingual architecture. *Frontiers in Psychology* 8. 2212.

Rothman, Jason. 2009. Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International Journal of Bilingualism* 13(2). 155–163.

Sasse, Hans-Jürgen. 1992. Language decay and contact-induced change: Similarities and differences. In Matthias Brenzinger (ed.), *Language death: Factual and theoretical explorations with special reference to East Africa*, 59–79. Berlin: Mouton de Gruyter.

Savedra, Monica Maria Guimaraes. 2020. Language vitality and transculturalization of european immigrant minorities: pomeranian in brazil. *Revista Diadorim* 22. 79–92.

Scontras, Gregory, Maria Polinsky & Zuzanna Fuchs. 2018. In support of representational economy: Agreement in heritage Spanish. *Glossa* 3(1). 1–29. DOI: 10.5334/gjgl.164.

Scontras, Gregory & Michael T. Putnam. 2020. Lesser-studied heritage languages: An appeal to the dyad. *Heritage Language Journal* 17(2). 152–155.

Shannon, Claude E. 1948a. A mathematical theory of communication. *Bell System Technical Journal*. 379–423.

Shannon, Claude E. 1948b. A mathematical theory of communication. *Bell System Technical Journal*. 623–656.

Sherkina-Lieber, Marina. 2015. Tense, aspect, and agreement in heritage Labrador Inuttitut: Do receptive bilinguals understand functional morphology? *Linguistic Approaches to Bilingualism* 5(1). 30–61.

Trudgill, Peter. 2011. *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford: Oxford University Press.

Wintai, Tsehaye, Tatiana Pashkova, Rosemarie Tracy & Shanley E. M. Allen. 2021. Deconstructing the native speaker: Further evidence from heritage speakers for why this horse should be dead! *Frontiers in Psychology* 12. DOI: 10.3389/fpsyg.2021.71735.

Chapter 2

Expressing diminutive meaning in heritage Twi: The role of complexity and language-specific preferences

Felix Kpogo[a], Alexandra Kohut[a] & Charles B. Chang[a]
[a]Boston University

Twi (Akan) and English can both express diminutive meaning using a morphological strategy (diminutive suffix) or a syntactic strategy (adjectival construction), but they differ with respect to native-speaker preferences – morphological in Twi, syntactic in English. Each strategy in Twi, moreover, is associated with different types of complexity (morphological, phonological, lexical, discourse-pragmatic, and/or inhibitory). In this study, we examined whether English-dominant, second-generation (G2) speakers of Twi in the US would express diminutive meaning in Twi differently from first-generation (G1) speakers. Results from elicited production suggest that G2 does indeed differ from G1 in this respect: whereas G1 relies on the morphological strategy, G2 relies on the syntactic strategy, producing adjectives post-nominally in accordance with Twi syntax. These results are discussed in light of variation in G2 speakers' morphological awareness and verbal fluency in Twi. Overall, our findings suggest that both the incremental complexity of linguistic options within a bilingual language repertoire and cross-linguistic influence at the level of preferences play a role in explaining G2's diminutive production.

1 Introduction

When presented with variable input, heritage speakers (HSs) tend to simplify complex forms in the heritage language (HL; see Kim 2007, Isurin & Ivanova-Sullivan 2008, Ivanova-Sullivan 2014, Scontras et al. 2015, 2017), resulting in systematic differences between their grammars and those of native speakers who

Felix Kpogo, Alexandra Kohut & Charles B. Chang. 2024. Expressing diminutive meaning in heritage Twi: The role of complexity and language-specific preferences. In Maria Polinsky & Michael T. Putnam (eds.), *Formal approaches to complexity in heritage languages*, 15–41. Berlin: Language Science Press. DOI: ??

continue to be dominant in the language.[1] Such differences, often observed in "interface" phenomena, have been attributed to the greater vulnerability of structures governed by an interface (e.g., the syntax-pragmatics interface) as compared with those situated within core domains of the grammar, which tend to be more resistant to simplification and reanalysis by HSs (Hulk & Müller 2000, Sorace 2000, Tsimpli et al. 2004).

In the current study, we examined the use of diminutives in the Twi of English-dominant HSs as a means of further investigating HSs' tendency towards simplification in the context of multiple linguistic options with different kinds of complexity.[2] The linguistic options of interest here are two strategies for expressing the notion of "smallness" in Twi: (1) a diminutive morpheme *-ba*/*-wa* (e.g., *sekan-ba* 'machete-DIM' → *sekamma* 'knife') and (2) an adjectival construction using the word *ketewa* 'little' (e.g., *sekan ketewa* 'knife', lit. 'machete little'). Crucially, these options are associated with different complexities for an English-dominant HS. On the one hand, the diminutive morpheme is complex in terms of transparency, allomorphy, and productivity (see §2.2). On the other hand, the adjectival construction is complex in terms of morphosyntactic conflict with English, given that adjectives are generally post-nominal in Twi but pre-nominal in English.

Because both of the above options for expressing diminutive meaning exist in the HL (Twi) and the majority language (English), our focus in this study was not on cross-linguistic differences in grammaticality, but rather on cross-linguistic differences in PREFERENCES for one option vs. the other. In particular, we examined whether second-generation (G2) Twi speakers in the US, under the influence of different linguistic preferences for English, would exhibit linguistic preferences for Twi that diverged from those of first-generation (G1) Twi speakers in the same environment. Under the assumption that they would, we also examined whether individual differences among G2 speakers in the strength of their observed preferences would be related to aspects of their HL proficiency – in particular, morphological awareness and verbal fluency.

[1]Following research arguing that HSs should be considered part of a continuum of native speakers (Rothman & Treffers-Daller 2014, Wiese et al. 2022), we use the term "heritage speaker" not in opposition with "native speaker", but rather in the sense of "switched-dominance bilingual" – that is, in contrast to the acquisition profile often assumed for native speakers who continue to be dominant in the target language (i.e., a profile that involves early, continuous, and socially robust exposure, but that may or may not be monolingual).

[2]The term "Twi" is used in this paper as a cover term to refer primarily to the Asante and Akuapem dialects of Akan. The term "Akan" is a generic name for at least eleven dialects, among which Fante, Asante, and Akuapem have achieved literary status. Asante and Akuapem, plus some other dialects, are often referred to collectively as Twi.

In the rest of the chapter, we report the results of a study designed to address the above questions. This study contributes to our understanding of innovative linguistic preferences in HSs; broadens the body of research on HLs by adding data on Twi (which remains understudied as a HL); and raises additional questions for future research on how HSs deal with different complexities of the HL.

2 Background

2.1 Heritage language grammars

Over the past two decades, the population of language users referred to as "heritage speakers" (HSs) has become the focus of a vibrant research program within the field of language acquisition. Although definitions of what constitutes a HS vary (see Polinsky 2006, Montrul 2008, Rothman 2009, Benmamoun et al. 2013b, Scontras et al. 2015), generally HSs are described as being exposed to rich, and often native-like, linguistic input in the HL from birth until some point before first language (L1) development is complete, when the primary source of linguistic input switches to a second language (L2; typically, the majority language of the society). This switch results in reduced contact with the HL and, ultimately, a change in language dominance (Polinsky 2008). Such an acquisition profile is common among second-generation immigrants, international adoptees, and members of multilingual societies, as well as other groups.

HSs are known to be a heterogeneous group, partly due to the fact that the quantity and quality of linguistic input in the HL after the point of reduced contact may vary considerably across individuals depending on factors such as access to formal education in the HL, the sociolinguistic status of the HL, the presence or absence of a HL speech community, and the age of reduced contact (Montrul 2010). Thus, HSs fall neither fully within the purview of L1 acquisition research nor fully within that of L2 acquisition research; rather, they comprise a spectrum of language users who exhibit a range of patterns. For instance, on a task testing clitic left dislocation and differential object marking in Spanish, English-dominant HSs of Spanish performed in between Spanish-dominant speakers and L2 learners, suggesting that HSs, while not quite like Spanish-dominant speakers, were also distinct from L2 learners (Montrul 2010).

Research on HSs has pointed toward two important considerations in the study of HL grammars. First, a tightly controlled experimental design is needed in order to differentiate among various possible outcomes in HL acquisition, which include dominant language transfer, interrupted acquisition leading to divergent attainment, and language attrition. These possible outcomes can give rise to

patterns in performance that overlap (Scontras et al. 2015), which may make it unclear how a given pattern should be interpreted. For instance, in a study of Brazilian Portuguese HSs, what appeared to be divergent attainment was actually language change in progress (Pires & Rothman 2009). Second, studies of HL grammars must account for attested sources of divergent patterns in a HL, including incipient changes in the HL input, resource constraints, and universal principles of language structure (Scontras et al. 2015). Universal principles are particularly relevant to the current study because research suggests that these principles guide the manner in which HSs tend to reduce complexity in the input, such as in loss of irregular morphology and reduction in morphology overall (Benmamoun et al. 2013a,b), movement toward less flexible word order (Isurin & Ivanova-Sullivan 2008, Ivanova-Sullivan 2014), and loss of non-compositional structure (Dubinina 2012, Rakhilina & Marushkina 2014).

In short, HSs exhibit a range of developmental profiles, which may render surface structures and production patterns ambiguous in terms of how they should be interpreted with respect to the HL system. Given this ambiguity, explanations for HSs' linguistic performance such as reduced complexity corresponding to universal principles cannot be ruled out without careful consideration. In this chapter, we focus on diminutive expression in Twi as a HL, which presents an interesting case of preexisting optionality between two grammatical forms that are associated with different complexities for English-dominant speakers. In the next section, we describe these complexities in more detail.

2.2 Diminutive expression in Twi and English

Cross-linguistically, diminutives are generally used to express the "smallness" of an entity. Twi and English are similar in that both can express diminutive meaning using a morphological strategy (i.e., diminutive suffix) or a syntactic strategy (i.e., adjectival construction), as shown in (1) and (2). However, the two languages differ in terms of native-speaker preferences for diminutive expression: in Twi, the morphological strategy is preferred, whereas in English, the syntactic strategy is preferred. Morphologically, Twi expresses the diminutive using the suffixes *-ba* and *-wa*,[3] and English using a variety of suffixes such as *-let*/*-lette*, *-y*/*-ie*, and *-ling*.

[3]There is dialectal variation in Akan in the realization of the suffix: Twi speakers employ both forms (e.g., *kuro-wa* 'small town', *anomaa-ba* 'baby/small bird'), whereas Fante speakers always use *-ba* (e.g., *dan-ba* 'cottage', cf. *dan* 'house').

(1) Diminutive expression in Twi (a: morphological, b: syntactic)

a. sekan-ba (> sekamma)
machete-DIM
'knife'

b. sekan ketewa
machete little
'knife'

(2) Diminutive expression in English (a: morphological, b: syntactic)

a. bagg-y

b. little bag

As in many other languages (e.g., Spanish: Marrero et al. 2007; Lithuanian: Savickiene 1998; Hebrew: Ravid 1998), the morphological diminutive (i.e., the diminutive exponed through a bound morpheme) in both Twi and English exists at the semantics-pragmatics/discourse interface, as exemplified in (3) and (4). The morphological diminutive in (3a) is ambiguous between a semantic meaning of smallness (which is available generally for this morpheme) or a discourse-pragmatic interpretation reflecting speaker attitudes (Appah & Amfo 2011); in contrast, the adjectival construction in (3b) is not ambiguous and can carry only the semantic meaning. Similarly, the morphological diminutive in (4a) can express either the semantic meaning of smallness or the pragmatics of speaker affection, but this does not hold for (4b). Thus, only (4a) may lead to infelicity in a context that is inconsistent with speaker affection.

(3) Possible interpretations of Twi diminutives

a. ade-wa
thing-DIM
'small thing'/'insignificant thing'

b. ade ketewa
thing little
'small thing'

(4) Possible interpretations of English diminutives

a. dogg-y (cf. #I despise the doggy from next door.)

b. little dog (cf. I despise the little dog from next door.)

The two strategies for diminutive expression in Twi are associated with different types of complexity for the Twi-English bilingual. Beginning with the morphological strategy, using the morphological diminutive appropriately requires navigating the discourse-pragmatic complexity of its dual meanings (semantic and pragmatic), described above. In addition, there is significant variation in the form of the morphological diminutive. Some of this variation comes from suppletive allomorphy conditioned on animacy: the diminutive suffix surfaces as *-ba* on animate stems and as *-wa* on inanimate stems (Appah & Amfo 2011). However, to make matters even more complex, these allomorphs are not strictly conditioned on animacy; rather, they are semi-lexically conditioned, as demonstrated by the allomorph for animates occurring with the inanimate stem in (1a). The diminutive suffix can also trigger bidirectional morphophonological changes to the stem and the suffix (Dolphyne 1988, Agyekum 2010, Appah & Amfo 2011), as in (1a). Furthermore, depending on the diminutivized item, the diminutive and/or the stem it combines with may not be isolable. In some cases, the diminutive and stem are easy to isolate (e.g., *dua-wa* 'chewing stick'; cf. *dua* 'tree/stick/log'), but in other cases the diminutive is lexicalized in the base of the word and cannot be isolated (e.g., *apakyiwa* 'small calabash with a cover'; cf. **apakyi*, not a word).

An additional dimension of complexity for the morphological strategy, both in Twi and in English, is restricted availability (and associated memory demands). In comparison to the syntactic strategy, which is universally available in both languages, the morphological strategy is less consistently available, as demonstrated in (5) and (6). In Twi, the morphological diminutive, though productive, is not permitted with some items (e.g., #*toa-wa* 'small bottle'). Even if the items not permitting the morphological diminutive form a natural class (synchronically or diachronically), such a class is not transparent to the average speaker; therefore, this set of items must effectively be lexically specified. In this respect, the morphological strategy involves an additional cognitive (memory) load.

(5) Diminutive morpheme restrictions in Twi

a. *nhoma-wa / *nhoma-ba
book-DIM
'little book'

b. nhoma ketewa
book little
'little book'

(6) Diminutive morpheme restrictions in English

a. *deer-y

b. little deer

Turning to the syntactic strategy, there is complexity here as well – namely, inhibitory complexity arising from a morphosyntactic conflict between Twi and English. Whereas adjective ordering is post-nominal (i.e., noun then adjective) in Twi, it is pre-nominal in English, creating a potential for cross-linguistic competition any time a Twi-English bilingual, particularly one dominant in English, uses an adjectival construction in Twi. Competition from a different possible word order within the bilingual language repertoire is known to influence the adjective ordering produced by early bilinguals (Nicoladis 2006). Therefore, given the conflicting word orders of Twi and English, use of the syntactic strategy in Twi may be associated with higher inhibition costs (related to suppressing the word order of English) than the morphological strategy for English-dominant HSs of Twi.

Crucially, this logic concerning the complexity of the syntactic strategy is based on the assumption that a conflict between different grammars in a bilingual repertoire introduces a cross-linguistic type of complexity (related to inhibition of a conflicting grammar) that is relevant for drawing predictions in this study. In regard to adjective ordering, we assume that if language *a* and language *α* implement adjectival modification of nouns syntactically differently, this conflict will make the task of adjectival modification in language *a* more complex (as compared to the case of language *α* patterning similarly to, and thereby reinforcing, language *a* or the case of there being no competing language *α* at all). From this assumption, it follows that a Twi HS's dominance in the English system (i.e., pre-nominal adjective syntax), which conflicts with the Twi system (i.e., post-nominal adjective syntax), will increase the complexity of using the Twi system. That said, this complexity may not necessarily pose much of a problem (see §2.3).

Since we have now introduced a cross-linguistic type of complexity into the discussion, it is worth considering the similarities and differences between Twi and English more broadly. In particular, might there be sources of competition from English that would make the inhibitory complexity of the morphological strategy even greater than that of the syntactic strategy? There are two reasons to believe that cross-linguistic competition from English is, in principle, a bigger issue for use of the syntactic strategy than the morphological strategy. First, there is a cross-linguistic conflict in morphosyntactic ordering for the adjectival construction only, as the diminutive morpheme is consistently suffixal and thus ordered the same with respect to the stem in both languages. Second, the

main dimensions of complexity for the morphological strategy – suppletive allomorphy, morphophonological alternations, lexical restrictions, and discourse-pragmatic ambiguity – mostly do not have straightforward correspondents in English to serve as competitors. To be specific, there is no allomorphy in English that resembles the *-ba*/*-wa* allomorphy of Twi formally or semantically, nor is there a phonological process of English that resembles the bidirectional assimilation evident in (1a). Furthermore, given the formal differences between Twi and English lexical items, it is difficult to link any lexical restriction of English to lexical restrictions of Twi (e.g., *sekan* 'machete' is not a phonological neighbor of any English word). With respect to discourse-pragmatic ambiguity, there is a cross-linguistic correspondence; however, the discourse-pragmatic interpretations of the diminutive suffix in Twi are generally paralleled – as opposed to contradicted – by the discourse-pragmatic interpretations that are possible for diminutive suffixes in English (e.g., affection, pejoration, etc.). This is not to say that English does not have suppletive allomorphy, phonological rules, lexical restrictions, or discourse-pragmatic ambiguities; rather, we argue that none of these aspects of English are likely to interfere with using the morphological strategy in the same way that the clearly reverse (vis-à-vis Twi) adjective ordering of English may intefere with using the syntactic strategy in a target-like fashion.

In short, both strategies for expressing smallness in Twi (morphological and syntactic) present complexity for the Twi-English HS bilingual. As such, whatever HSs' preferences may be for one strategy over another, they cannot be explained straightforwardly in terms of eliminating complexity as any choice will result in trading, as opposed to eliminating, complexities. Further, because both strategies are amply available in Twi, HSs' preference for one particular strategy cannot be explained by a lack of exposure to the other strategy. This all leads to the central question in this study: which strategy do Twi HSs prefer, the morphological strategy (diminutive suffix) or the syntactic strategy (adjectival construction)?

2.3 Research questions and hypotheses

The current study addressed three research questions:

1. Do English-dominant, second-generation (G2) Twi speakers in the US differ from first-generation (G1) Twi speakers with respect to preferred strategy for expressing the semantic notion of smallness?

2. If G2 differs from G1, does the difference between groups reflect a simplification in available strategies for G2 or a more subtle shift in G2's preferences for alternative strategies with different complexities?

3. Do individual-difference variables for Twi predict rates of morphological diminutive use in elicited production by G2?

We had a specific hypothesis in respect to each of the above three questions. Our hypothesis concerning question 1 (H1) was that G2 would prefer the syntactic strategy for expressing smallness over the morphological one (and, thus, would differ from G1 in terms of preferred strategy). The logic behind H1 is that the main complexity associated with the syntactic strategy is one that must be dealt with not just in expressing smallness, but in basic use of the HL, because all adjectives in Twi are post-nominal. That is, the INCREMENTAL complexity of the syntactic strategy over and above what must be mastered for basic use of the HL is nil. By contrast, the incremental complexity of the morphological strategy is considerable, as the meaning variation, form variation, and lexical restrictions connected with this strategy probably go beyond what must be mastered for basic use of the HL. Therefore, under the assumption that English-dominant G2 speakers who are capable of basic use of the HL are generally motivated to minimize complexity when using the HL, G2 should tend toward the syntactic strategy.

Our hypothesis concerning question 2 (H2) was that G2's preference for the syntactic strategy would be clear but not categorical, consistent with a shift in preferences rather than simplification of mental representations related to the diminutive. The logic behind H2 is based partly on previous evidence of HSs' behavior approximating, if not replicating, that of target language-dominant speakers (e.g., Montrul 2010) and partly on the Interface Vulnerability Hypothesis (Sorace & Filiaci 2006, Sorace 2011), which suggests that interface phenomena are particularly vulnerable in the context of bilingualism. Since we argue that the morphological strategy is in fact an interface phenomenon, it follows that it will be vulnerable to interrupted acquisition, cross-linguistic influence (CLI), and/or attrition. Thus, we expected the morphological strategy to have been acquired to some degree, but to be weaker – and thus less preferred – than the syntactic one.

Given H2, our hypothesis concerning question 3 (H3) was that individual differences in relevant aspects of Twi proficiency (in particular, verbal fluency and morphological awareness) would, indeed, help predict rates of morphological diminutive use. Because we expected that G2 as a group would not categorically reject the morphological strategy (H2), this leaves room for variation in G2's use of the morphological strategy, and we predicted that this variation would be related to Twi proficiency. More specifically, we predicted that higher verbal fluency and higher morphological awareness would be associated with higher rates of morphological diminutive use (i.e., the preferred strategy for G1).

3 Methodology

3.1 Participants

Speakers of the Asante Twi dialect of Akan, participants comprised two groups residing in the same region of the US: a group of second-generation (G2) speakers (N = 19; 16f, 3m; M_{age} = 21.7 yr, range 13–33) and a control group of first-generation (G1) speakers (N = 8; 4f, 4m; M_{age} = 42.6 yr, range 21–74). Another six G2 speakers were tested but ultimately excluded from analysis because they did not meet the minimum age requirement (age 12) and/or exceeded the upper limit for age of arrival in the US allowed for this group (age 5).

Members of the G2 group were all early arrivals to the US. The majority (15/19) were born in North America (the US or Canada), while four immigrated to the US at age 5 or younger. All were born to Twi-speaking parents, received input in Asante Twi at home starting in infancy, and spoke Twi with their parents; however, according to questionnaire data, all were dominant in English.

In contrast, members of the G1 group were all late arrivals to the US, with an average age of arrival of 32.3 yr (range 18–52). Their average length of residence in the US was 9 yr (range 2–16) at the time of testing. All spoke Twi from birth as their primary language and reported using exclusively Asante Twi at home, with friends and local relatives, and at church. According to questionnaire data, all were dominant in Twi.

3.2 Procedure

The study was carried out via a combination of virtual and in-person testing sessions. Data from G1 were collected virtually (via Zoom), while data from G2 were collected virtually (n = 11) and in person (n = 8). The basic protocol and format of the tasks in the task battery, as well as compensation, were the same in virtual and in-person testing sessions. The testing sessions were conducted primarily in Asante Twi by the first author, a native speaker. Occasionally, when a G2 participant could not understand or recall a word, the experimenter would briefly switch to English to accommodate the participant and ensure their understanding (such language switches did not affect rates of morphological diminutive use; see §4.4). All tasks involving oral responses were audio-recorded.

The task battery consisted of four tasks: a picture description task, an acceptability judgment task, a morphological parsing task, and a verbal fluency task, each described further in §3.3. Participants also completed a detailed language background questionnaire. All tasks were completed by both groups except for

the English morphological parsing task, which was completed by G2 only. The full questionnaire and materials used in all tasks are publicly available on the Open Science Framework (OSF) at https://osf.io/cze5g/.

To minimize priming effects across tasks, we incorporated two constraints on task order within each testing session. First, the picture description task was administered before the acceptability judgment task and the morphological parsing task. Second, the Twi morphological parsing task was administered before the English one. The verbal fluency task and the background questionnaire were not strictly ordered with respect to the other tasks.

3.3 Tasks

3.3.1 Picture description task

The goal of the picture description task was to examine participants' use of the morphological and syntactic strategies for expressing smallness. In this task, participants were shown a series of slides, each consisting of four pictures: a target picture, two related pictures which differed from the target picture in terms of a specific attribute (these served as standards for comparison), and an unrelated distractor picture. For each slide (trial), participants were told (in Twi), "Here are four things. Kofi wants this thing," at which point the target picture was circled on the slide. Participants were then asked what Kofi wanted and gave an oral response to indicate the target item. At the end of each trial, participants were asked if they could think of any other way they might describe the target item.

There were 25 slides total, which were presented in the same pseudo-random order to all participants. On ten slides, the target picture (of a relatively small item) was meant to elicit a diminutive form, either a noun suffixed with the morphological diminutive or a noun phrase (NP) modified by an adjective. On three slides, the target picture (of a control item) was meant to elicit an unmodified noun rather than a diminutive form. On twelve slides, the target picture (of an unrelated distractor item) was meant to elicit various non-diminutive forms, such as spatial expressions, plural forms, and NPs modified by emotion or color adjectives. Two of the slides in this last category were used in a short practice session to familiarize participants with the task before beginning the test trials.

3.3.2 Acceptability judgment task

An acceptability judgement task was administered to test whether the morphological diminutive was available for the target items within participants' grammars. The stimuli in this task consisted of Twi noun+adjective phrases and were

presented in a survey administered through Qualtrics (Qualtrics 2021). On each trial, participants saw one of the phrases written on screen and played an audio file of a Twi native speaker pronouncing the phrase (which could be replayed an unlimited number of times). Participants then provided an acceptability judgment indicating whether or not the phrase was something they could say in Twi (or, alternatively, a third option, "I don't know"). The task was untimed and self-paced.

The stimulus set for this task included 30 test items: 10 critical items, 10 control items, and 10 filler items. The critical items consisted of the forms with the morphological diminutive used in the picture description task. The control items consisted of five grammatical and five ungrammatical items unrelated to the diminutive. The ungrammatical control items contained errors in plural marking, errors in adjective placement, and semantic ill-formedness. Two additional control items (one grammatical and one ungrammatical) were used in a short practice session to familiarize participants with the task before the test trials. As for the filler items, these were parallel to the critical and control items in terms of form, consisting of noun+adjective collocations. However, in contrast to the control items, the filler items contained plural markers that are subject to idiolectal morphological variation, both suppletive and phonological, which is not currently well understood. The filler items comprised six items with suppletive variation and four items with phonological variation.

3.3.3 Morphological parsing task

A morphological parsing task in English and in Twi was included in the task battery for two reasons. First, production of a morphological diminutive in the picture description task may not necessarily reflect a form that was morphologically complex for the speaker; therefore, data from morphological parsing offered insight into whether forms with the morphological diminutive were mentally represented by participants as complex (i.e., stem+suffix) or simplex (i.e., lexicalized). Second, given previous findings of an association of morphological awareness across languages, at least for bilinguals who acquire typologically distant languages (Hayashi & Murphy 2013), we wanted to measure G2's morphological awareness in both of their languages so as to consider a broad index of morphological awareness as a predictor of performance in the picture description task.

Like the acceptability judgement task, the morphological parsing task was administered in a survey format through Qualtrics, with separate surveys for Twi and English. The stimuli consisted of words presented in isolation. On each

trial, the target word was presented orthographically with sequentially numbered lines between each pair of consecutive letters, and auditorily via an audio file of a Twi native speaker pronouncing the word (which could be played an unlimited number of times). Participants were instructed to select the number(s) of the line(s) dividing the word into its meaningful parts (or, alternatively, the option "there is nowhere to divide this word"). The task was untimed and self-paced.

The stimulus set for each of Twi and English consisted of 32 test items (including 10 monomorphemic items). For each language, an additional morphologically complex item and monomorphemic item were used in a short practice session to familiarize participants with the task before the test trials.

3.3.4 Verbal fluency task

A verbal fluency task was used to measure participants' Twi proficiency, which was later considered as a predictor in the analysis of the data from the picture description task. In this task, participants were asked to name as many items within a given semantic domain as they could in 60 seconds. Each participant did this for two semantic domains, one being food and the other being either animals or environment/habitat (randomly assigned).

4 Results

Before presenting the results of the picture description task, we summarize the results of the acceptability judgment task, the morphological parsing task, and the verbal fluency task. Data from the latter two tasks are incorporated into the analysis of the picture description results as predictor variables. All analyses were completed in R (R Core Team 2021) using the `lme4` and `optimx` packages (Nash & Varadhan 2011, Nash 2014, Bates et al. 2015). The full dataset is available on the OSF at https://osf.io/k6spv/.

4.1 Acceptability judgment results

The analysis of the acceptability judgment task focused on responses to the 20 critical and control items. Responses from one G2 participant were excluded because they uniformly accepted all of the items, making it unclear whether they understood the task. The dataset consisted of 520 total responses (= 20 items x 26 participants), of which there were only 36 "uncertain" (i.e., "I don't know") responses (two from G1, 34 from G2). Although one G2 participant gave "uncertain"

responses to seven items, all others did so to four or fewer items. With respect to individual items, there were 18 items which received "uncertain" responses; the most "uncertain" responses received by a single item, which was a critical item, was six, and all other items received four or fewer such responses. Given the generally low number of "uncertain" responses at both the participant level and the item level, these were excluded from further analysis (rather than being grouped with either the "acceptable" or "unacceptable" responses). Thus, the final dataset submitted to statistical analysis consisted of 484 responses (i.e., 93.1% of the responses to critical and control items).

The likelihood of accepting items in this task was analyzed in terms of a mixed-effects logistic regression model (Model 1), which focused on two comparisons: the comparison between grammatical and ungrammatical control items and that between grammatical control and critical items. Model 1 included Group (reference level = G1) and ItemType (reference level = grammatical control) as treatment-coded fixed effects and random intercepts by Participant and Item.

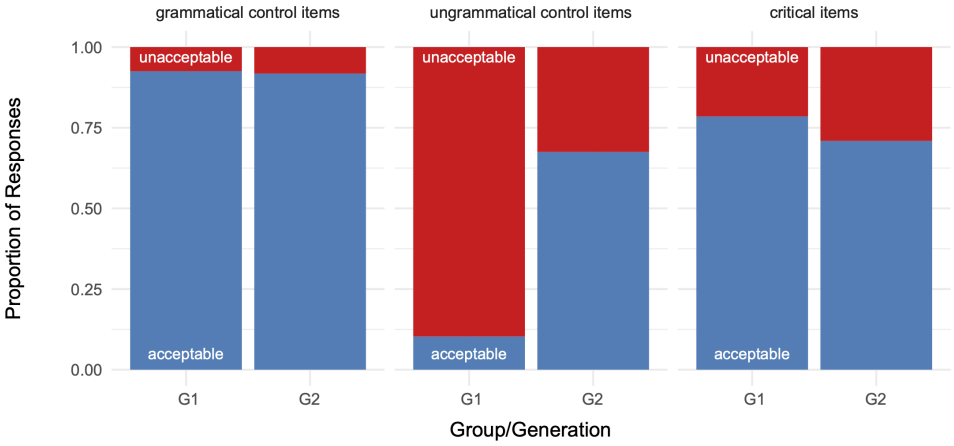


Figure 1: Distribution of responses in the acceptability judgement task by item type (grammatical control, ungrammatical control, critical) and group (G1, G2). Blue and red indicate, respectively, responses accepting and responses rejecting the test items.

Starting with the first comparison, as shown in Figure 1, both G1 and G2 were more likely to accept than reject grammatical control items; however, the similarity between G1 and G2 on grammatical control items was not seen in ungrammatical control items, which G2 was much more likely to accept than G1 was (consistent with the "yes-bias" documented for HSs; Polinsky 2018). The group disparity in accepting ungrammatical control items specifically was re-

flected in an analysis of variance (ANOVA) on Model 1 using the `Anova()` function in the `car` package (Fox & Weisberg 2019). The ANOVA revealed no main effect of Group [$\chi^2(1) = 0.783, p = 0.376$] but a significant main effect of ItemType [$\chi^2(1) = 19.648, p < 0.001$] and a significant Group × ItemType interaction [$\chi^2(1) = 27.428, p < 0.001$].

Table 1: Fixed effects in Model 1 of the likelihood of accepting test items (grammatical control, ungrammatical control, critical) in the acceptability judgement task [$N = 484$, log-likelihood = -221.1]. Intercept represents Group = G1, ItemType = grammatical control. Significance code: *** $p < 0.001$.

Predictor	β	SE	z	Pr(> |z|)
(Intercept)	2.947	0.786	3.748	< 0.001 ***
Group: G2	−0.003	0.887	−0.003	0.997
ItemType: ungrammatical	−5.483	0.962	−5.699	< 0.001 ***
ItemType: critical	−1.334	0.755	−1.768	0.077
Group: G2 × ItemType:				
ungrammatical	3.540	1.013	3.493	< 0.001 ***
critical	−0.477	0.830	−0.575	0.566

The fixed-effects coefficients of Model 1 are summarized in Table 1. The results of Model 1 indicated that G1 was significantly more likely to accept grammatical control items compared to the null hypothesis, i.e. 50-50 odds [$\beta = 2.947$, $p < 0.001$]; crucially, however, G1 was also significantly less likely to accept ungrammatical than grammatical control items [$\beta = -5.483, p < 0.001$]. G2 did not significantly differ from G1 in terms of likelihood of accepting grammatical control items [$\beta = -0.003, p = 0.997$]. However, for G2, the reduction in likelihood of accepting ungrammatical control items compared to grammatical ones was significantly smaller than seen in G1 [$\beta = 3.540, p < 0.001$], suggesting that G2 did not reject ungrammatical control items as readily as G1 did.

Turning to the second comparison, as shown in Figure 1, both G1 and G2 were more likely to accept than reject critical items, much like grammatical control items, although in general participants were less likely to accept critical items than grammatical control items. Crucially, the results of Model 1 indicated little difference between G1 and G2 in this respect. G1 was not significantly more or less likely to accept critical items compared to grammatical control items [$\beta = -1.334, p = 0.077$]. Furthermore, the `Group:G2 × ItemType:critical` interaction coefficient was negative but not significant [$\beta = -0.477, p = 0.566$],

meaning that the reduction in likelihood of accepting critical items compared to grammatical control items was statistically similar for G2 relative to G1. In short, the general similarity between G2 and G1 on grammatical control items was also reflected in critical items containing the Twi diminutive suffix (crucially, the same items elicited in the picture description task), meaning that any between-group disparity in use of the morphological strategy is unlikely to be due to group differences in the suffix's acceptability per se.

4.2 Morphological parsing results

Responses in the Twi morphological parsing task were analyzed to calculate a Twi morphological awareness score for each participant. Participants received a point for every target morpheme boundary they identified, and these points were then totaled. Participants' raw point totals were then converted to z-scores by group. For G2, who also completed an English morphological parsing task, we used the same method to calculate individual English morphological awareness scores as well; however, ultimately only the Twi morphological awareness scores were used as a predictor in the model of the picture description results. Although raw Twi morphological awareness scores were, on average, slightly higher for G1 ($M = 23.4, SD = 4.3$) than G2 ($M = 19.9, SD = 4.0$), this difference was not significant [Welch-corrected two-sample $t(12.5) = 1.967$, $p = 0.072$].

To get a picture of whether participants analyzed the Twi diminutive suffix as a separate morpheme, we also calculated the percentage of target diminutive morpheme boundaries identified by each group. This analysis showed that the majority of diminutive morpheme boundaries were successfully identified by G1 ($M = 79.1\%, SD = 23.1$) and by G2 ($M = 71.3\%, SD = 18.4$), suggesting that both groups represented the Twi diminutive suffix as a distinct unit in the Twi grammar (and not as merely lexicalized in the words in which the suffix occurs). Thus, this result supports interpreting any disproportionate use of the syntactic strategy for expressing diminutive meaning in Twi by G2 as representing an innovative preference rather than the loss of the diminutive morpheme (see research question 2 in §2.3).

4.3 Verbal fluency results

Raw scores in the verbal fluency task were tabulated as the number of items named by participants for a target domain. Because each participant was assigned two domains, the second of which varied across participants, the raw scores from the two options for the second domain (habitat, animals) were compared statistically to check for a domain effect on naming performance. This

comparison showed that, although scores tended to be higher on the "animals" domain ($M = 8.6, SD = 6.1$) than the "habitat" domain ($M = 6.8, SD = 2.4$), the difference between these domains was not significant [Welch-corrected two-sample $t(15.5) = 1.009$, $p = 0.328$]. Consequently, for the purposes of generating by-participant verbal fluency scores to use as a predictor in the model of the picture description results, we calculated a composite score for each participant by summing the number of unique items named across both domains. As expected, the raw composite scores were significantly higher for G1 than G2 [G1: $M = 28.9, SD = 7.2$; G2: $M = 17.8, SD = 5.5$; $t(10.6) = 3.885$, $p = 0.003$]. The raw composite scores were subsequently converted to z-scores by group, and the standardized scores were used in modeling.

4.4 Picture description results

To examine participants' likelihood of using the morphological strategy for talking about smallness, responses in the picture description task were coded in binary fashion as either "morphological" or "non-morphological". Responses in the "morphological" category included responses with the diminutive suffix only (e.g., *sekam-ma* 'knife'), responses where an adjectival construction was produced initially and then a suffixed form was produced (e.g., *sekan ketewa* 'knife' and then *sekam-ma* 'knife'), and responses where a simplex form was produced initially and then a suffixed form was produced (e.g., *sekan* 'machete' and then *sekam-ma* 'knife'). Responses in the "non-morphological" category included responses with an adjectival construction only, responses with a simplex form only, and responses where a simplex form was produced initially and then an adjectival construction was produced. Responses that included a diminutive suffix and an adjective within the same form ($N = 4$, all produced by G1; 1% of all responses) were considered ambiguous in terms of preference for the morphological strategy and were therefore excluded from analysis.

The distribution of responses across categories is shown in Figure 2, which separates out "syntactic + morphological" responses (i.e., those where an adjectival construction was produced before the final suffixed form was produced) and excludes the very few combined suffix-with-adjective responses for clarity. As shown in Figure 2, G1 and G2 differed markedly from each other in terms of preferences for expressing diminutive meaning, G1 preferring the morphological strategy and G2 the syntactic strategy. To analyze participants' response data statistically, we built two additional mixed-effects logistic regression models on the likelihood of producing the morphological diminutive (Models 2 and 3). Model 2 focused on the aforementioned group difference (see research question 1 in

§2.3) and contained the fixed effect Group (treatment-coded; reference level = G1). Model 3 focused on the efficacy of individual-difference variables for predicting individual variation, particularly in G2 (research question 3); thus, Model 3 contained fixed effects for Twi morphological awareness score (MorphAwareness) and Twi verbal fluency score (Fluency), both standardized by group. Both models had a random-effects structure consisting of random intercepts by Participant and by Item.

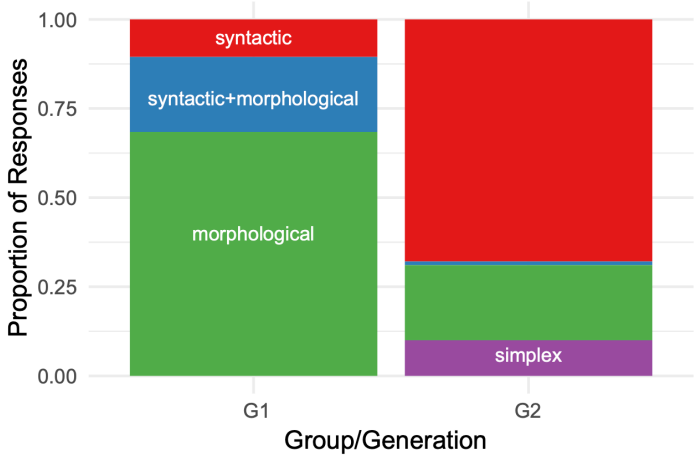


Figure 2: Distribution of responses in the picture description task by group (G1, G2). The two main categories of response are "morphological" (i.e., diminutive suffix) and "non-morphological" (e.g., syntactic: adjectival construction). Responses marked as "syntactic + morphological" (initial use of an adjectival construction, final use of a diminutive suffix) were grouped into the "morphological" category for analysis. Responses marked as "simplex" (no diminutive) were grouped into the "non-morphological" category for analysis.

Results of Model 2 confirmed that the group difference evident in Figure 2 was statistically significant. In particular, G2 was much less likely to produce the morphological diminutive than G1 was [$\beta = -6.366, z = -5.869, p < 0.001$].[4]

[4]As mentioned in §3.2, the experimenter sometimes switched into English to facilitate the testing procedure with G2. To explore whether such language switches may have affected G2's behavior, we conducted a post hoc analysis by pseudo-randomly selecting half of the critical trials including all G2 participants, transcribing the speech produced in these trials, and counting the number of instances of the experimenter switching into English. This analysis revealed that, in this set of 100 trials, the experimenter switched to English 19% of the time. However, rates of morphological strategy use were identical between trials with and trials without a switch (switch: 4/19 = 21%; no-switch: 17/81 = 21%), suggesting that the experimenter's switching into English did not play a significant role in G2's observed strategy use.

Turning to Model 3, we found evidence of an effect of verbal fluency but not of morphological awareness. The fixed-effect coefficients of Model 3 (summarized in Table 2) indicated that, at average levels of MorphAwareness and Fluency (i.e., z-score of zero), the odds of G2 producing the morphological diminutive were significantly lower than 50–50 [$\beta = -1.987$, $p = 0.035$]. At average levels of Fluency, higher levels of MorphAwareness were not associated with a significantly higher likelihood of producing the morphological diminutive [$\beta = 0.020$, $p = 0.941$]. On the other hand, at average levels of MorphAwareness, higher levels of Fluency were associated with a significantly higher likelihood of producing the morphological diminutive [$\beta = 0.992$, $p = 0.026$], and the interaction coefficient did not indicate a significant change in this effect at higher levels of MorphAwareness [$\beta = -0.746$, $p = 0.068$].

Table 2: Fixed effects in Model 3 of the likelihood of G2 producing a morphological diminutive in the picture description task [$N = 190$, log-likelihood $= -64.2$]. Significance code: * $p < 0.05$.

Predictor	β	SE	z	Pr(> |z|)	
(Intercept)	−1.987	0.941	−2.111	0.035	*
MorphAwareness	0.020	0.276	0.074	0.941	
Fluency	0.992	0.445	2.227	0.026	*
MorphAwareness × Fluency	−0.746	0.409	−1.822	0.068	

As a final part of the analysis of individual-difference variables, we inspected the omnibus correlation of each of MorphAwareness and Fluency with individual G2 participants' overall rate (proportion) of morphological diminutive production. As shown in Figure 3 (and consistent with Figure 2 showing the group pattern), at an individual level G2 did not show particularly high rates of morphological diminutive production; crucially, however, these rates were almost all higher than zero, meaning that few G2 participants showed evidence of loss of the diminutive suffix. As for MorphAwareness and Fluency, the correlation analyses showed that these variables were not significantly correlated with each other for G2 [Pearson's $R = 0.383, t(17) = 1.707, p = 0.106$]. Furthermore, morphological diminutive production was not significantly correlated with MorphAwareness [Pearson's $R = 0.263, t(17) = 1.122, p = 0.277$], but was significantly, and moderately, correlated with Fluency [Pearson's $R = 0.476, t(17) = 2.233, p = 0.039$].

Taken together, the results of Model 3 and the correlation analyses point to verbal fluency as a stronger predictor of G2's morphological diminutive production

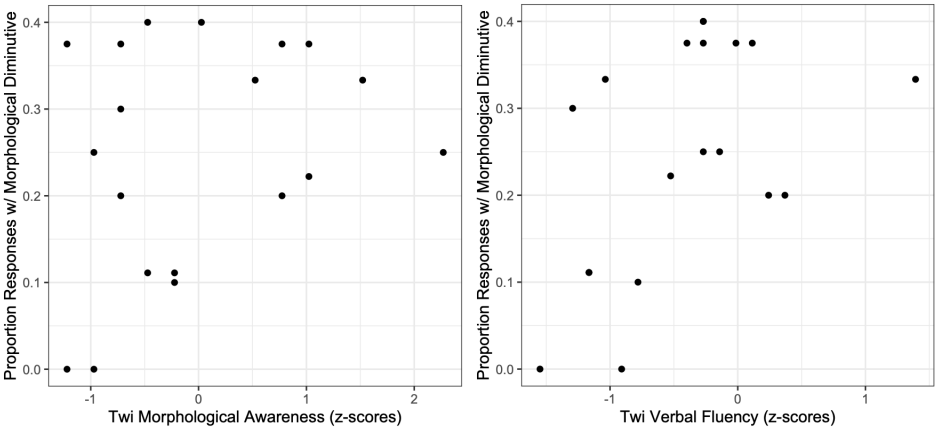


Figure 3: Scatterplots of G2's proportions of morphological diminutive production by Twi morphological awareness score (left) and verbal fluency score (right). Each dot represents one G2 participant.

than morphological awareness. However, we regard this finding with caution, as we observed rather high levels of morphological awareness among the G2 participants in this study overall (see §4.2). Thus, it is possible that the predictive power of morphological awareness in Twi may differ with a G2 sample evincing a wider range of morphological awareness.

5 Discussion

Returning to our hypotheses outlined in §2.3, recall that the current study tested three hypotheses (H1–H3) about English-dominant, second-generation (G2) Twi speakers' knowledge of the diminutive suffix (i.e., the morphological diminutive) and their relative preferences for the morphological and syntactic strategies of expressing diminutive meaning in Twi. Combining results from four tasks (acceptability judgment, morphological parsing, verbal fluency, and picture description), the findings of this study generally provided support for H1–H3. We consider each hypothesis in turn below.

First, we hypothesized that, whereas first-generation (G1) Twi speakers would show a preference for the morphological strategy of expressing diminutive meaning, G2 would show a preference for the syntactic strategy (H1). Results of the picture description task were consistent with H1: G1 strongly preferred the morphological strategy, using the diminutive suffix well over half of the time, but G2 consistently preferred the syntactic strategy, using the diminutive suffix less

than half of the time at both the group level and the individual level (see Figure 3). Thus, English-dominant G2 Twi speakers in the US do indeed show a different pattern with respect to strategies for expressing the semantic notion of smallness as compared to adult G1 speakers. Because the morphological strategy is incrementally more complex than the syntactic strategy (i.e., relative to the complexity that G2 must master for basic use of the HL apart from the diminutive), this finding is superficially consistent with the tendency of HSs to simplify complex linguistic phenomena in the HL.

Crucially, however, G2's bias toward the syntactic strategy does not reflect their having failed to acquire the morphological diminutive. On the contrary, despite the complexities associated with the morphological diminutive, the vast majority of G2 participants produced the morphological diminutive at least part of the time they needed to express diminutive meaning in the picture description task, suggesting they have not simplified the HL grammar by eliminating the morphological strategy entirely. This finding thus supports our hypothesis that G2's preference for the syntactic strategy, while stronger than G1's, is not categorical (H2). Results from the acceptability judgment and morphological parsing tasks further suggest that G2 generally represents the morphological diminutive in the HL grammar. First, G2 did not perform significantly differently from G1 on critical items with the diminutive in the acceptability judgement task. Second, G2 parsed the diminutive suffix as a meaningful unit in the Twi morphological parsing task. These results are inconsistent with a scenario in which G2 speakers have not acquired the morphological diminutive.

Turning to variation in G2, we found partial support for our hypothesis that individual differences in Twi morphological awareness and Twi verbal fluency would predict variation in G2's rates of morphological diminutive use (H3). In particular, we found an effect of verbal fluency, but not of morphological awareness: G2 participants with higher verbal fluency scores were more likely to produce a morphological diminutive in the picture description task. However, we also observed that G2's morphological awareness scores in Twi were high overall – in fact, not significantly different from G1's – leaving open the possibility of observing an effect of morphological awareness with a wider range in morphological awareness. Thus, further study of the role of morphological awareness in morphological diminutive use would be a useful direction for future research.

In connection with G2's observed preference for the syntactic strategy, it is important to note that virtually all of the G2 participants in this study had indeed acquired the post-nominal adjective syntax associated with the syntactic strategy. Because adjective order is thought to be an early-acquired aspect of core syntax (mastered as early as age 2; see Brown 1973, Paradis et al. 2000, Nicoladis

2002) and is consistently post-nominal throughout the Twi language (i.e., not just for the purposes of expressing the diminutive), we expected G2 to successfully manage the inhibitory complexity of using the syntactic strategy with target-like post-nominal adjective order (and not to transfer the competing pre-nominal adjective order of English), even if their preference for the syntactic strategy itself might be English-influenced. In accordance with this expectation, almost all the G2 participants consistently produced the post-nominal adjective order in Twi; further, of the two G2 participants who did not, only one showed clear evidence of transferring the English order to Twi, producing the English order in nearly all target items. Crucially, the overwhelmingly target-like production of Twi adjective order is inconsistent with the idea of unconstrained dominant language transfer to the weaker language, as has been suggested in other bilingual studies (Yip & Matthews 2000). Under the assumption that adjective order is part of core syntax, this finding instead supports the idea that early-acquired, core areas of the grammar in the weaker language remain stable over time and resistant to CLI (Sorace & Filiaci 2006, Sorace 2011), although there may be the occasional exceptional case as we observed in this study.

We are still left to explain what exactly is responsible for the observed divergence in linguistic preferences between G2 and G1, and we end this section by discussing the possible contributions of complexity, CLI, and universal tendencies. To begin, we believe that complexity – in particular, minimization of complexity – plays a role. As discussed in §2.3, for a variety of reasons the morphological strategy of expressing diminutive meaning in Twi can be considered incrementally more complex than the syntactic strategy for English-dominant speakers (i.e., G2). Consequently, the finding of a strong preference in G2 for the syntactic strategy – or, to put it another way, G2's move away from the morphological strategy preferred by G1 – is consistent with a tendency for HSs to minimize the complexity of using their HL. The operative word here is "minimize", as opposed to "eliminate" or "simplify" complexity, however, because it bears repeating that G2 still uses the morphological strategy, just less often than G1 does; that is, complexity minimization underlies the preferences, but not the availability of the strategies themselves. Converging evidence of complexity minimization comes from other aspects of G2's responses in the picture description task as well. For example, whereas G1 consistently distinguished *sekamma* 'knife' (a diminutivized form) and *sekan* 'machete' (a simplex form), and *adɔmma* 'little bell' (diminutivized) and *ɛdɔn* 'bell' (simplex), some G2 participants did not do so consistently, producing simplex forms in critical trials that called for the diminutive (see Figure 2). This type of response minimizes complexity by conflating lexical distinctions and avoiding phonological rules that apply with the diminu-

tive, although it does not necessarily indicate that the speaker would never make these distinctions or apply these rules.

CLI can also account for the divergent preferences of G2 vis-à-vis G1, and we believe that it plays a role as well. Because the syntactic strategy is preferred in the dominant language, English, G2's preference for the syntactic strategy in Twi can be interpreted as reflecting CLI from English preferences (or, to put it another way, from the relative strength of the syntactic strategy in English, where it is generally preferred over the morphological strategy). But, as above, if G2's preference for the syntactic strategy can be explained in terms of complexity minimization, why posit that CLI is involved at all? There is one aspect of our results that points to this conclusion. As mentioned above, we found that there were two G2 participants who, unlike other G2 participants, used pre-nominal adjective order to implement the syntactic strategy at least part of the time. Because this pre-nominal adjective order is ostensibly due to CLI from English, CLI must be invoked to explain these participants' production. Therefore, seeing no reason to believe that CLI is limited to adjective order, we assume that CLI also plays a role in the use of the syntactic strategy itself; it is just that, for most G2 speakers, this CLI is not allowed to extend into the core syntax of the HL.

As for universal tendencies, we do not consider it likely that a universal tendency is responsible for G2's preference for the syntactic strategy, because there is no clear typological or developmental evidence for a universal tendency that would favor the syntactic strategy. In regard to typology, we expect such a tendency to be reflected in a bias toward analytic languages (e.g., diachronic changes resulting in synthetic languages becoming more analytic, but not the other way around). Further, in first language development, such a tendency should produce a bias toward analytical constructions (e.g., two-word phrases emerging before bimorphemic words). To our knowledge, however, neither of these hypothetical biases is strongly supported by the literature; this includes the recent literature on creoles, which suggests that "creoles are not more analytic than the other [lexifier] varieties" (Siegel et al. 2014: 49). Thus, we conclude that G2's preference for the syntactic strategy is not due to a universal tendency, but instead attributable to complexity minimization and CLI.

6 Conclusion

In closing, we would like to acknowledge two limitations of this study, which point out directions for future research on the web of factors involved in shaping HSs' linguistic preferences in their HL. First, our sample of G2 Twi speak-

ers (HSs) was relatively small and possibly showed unusually high morphological awareness. Second, the task we used to measure diminutive production ultimately focused on elicited speech, which may not reflect how HSs express diminutive meaning in naturalistic speech communication. Thus, it would be useful in future work to replicate and extend the current findings with a larger, socio-demographically more diverse participant sample and with a task paradigm that more closely mimics spontaneous conversational speech.

Finally, another direction for future research is to begin to tease apart the effects of complexity and CLI, which are confounded in the current findings. In our case of Twi-English bilingualism, complexity minimization and CLI from English both favor the syntactic strategy of expressing diminutive meaning, so it is impossible to know for sure what the relative contribution of each factor is to G2's observed preference for the syntactic strategy. This type of question could be addressed by examining other cases of preferences in HS bilingualism, where complexity minimization favors one option but CLI from the dominant majority language favors a different option. Research in this vein would improve our understanding of the unique role of complexity in influencing HSs' use of their HL.

Acknowledgments

The authors gratefully acknowledge funding from the Department of Linguistics at Boston University, support from the Seventh-Day Adventist Church in Massachusetts and New York in reaching the Twi community, and helpful feedback from anonymous reviewers and audience members at the Bilingualism, Mind, and Brain Lab meeting and the 47th Boston University Conference on Language Development (BUCLD 47).

References

Agyekum, Kofi. 2010. *Akan kasa nhyehyeɛ*. Accra, Ghana: Dwumfour Publications.

Appah, Clement K. I. & Nana Aba Appiah Amfo. 2011. The morphopragmatics of diminutive morpheme (*-ba*/*-wa*) in Akan. *Lexis* 6. 85–103.

Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48.

Benmamoun, Elabbas, Silvina Montrul & Maria Polinsky. 2013a. Defining an "ideal" heritage speaker: Theoretical and methodological challenges reply to peer commentaries. *Theoretical Linguistics* 39(3–4). 259–294.

Benmamoun, Elabbas, Silvina Montrul & Maria Polinsky. 2013b. Heritage languages and their speakers: Opportunities and challenges for linguistics. *Theoretical Linguistics* 39(3-4). 129–181.

Brown, Roger. 1973. *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Dolphyne, Florence Abena. 1988. *The Akan (Twi-Fante) language: Its sound systems and tonal structure*. Accra, Ghana: Ghana Universities Press.

Dubinina, Irina. 2012. *How to ask for a favor: An exploration of speech act pragmatics in heritage Russian*. Bryn Mawr, PA: Bryn Mawr College. (Doctoral dissertation).

Fox, John & Sanford Weisberg. 2019. *An R companion to applied regression*. 3rd edn. Thousand Oaks, CA: Sage Publications.

Hayashi, Yuko & Victoria A. Murphy. 2013. On the nature of morphological awareness in Japanese-English bilingual children: A cross-linguistic perspective. *Bilingualism: Language and Cognition* 16(1). 49–67.

Hulk, Aafke & Natascha Müller. 2000. Bilingual first language acquisition at the interface between syntax and pragmatics. *Bilingualism: Language and Cognition* 3(3). 227–244. DOI: 10.1017/S1366728900000353.

Isurin, Ludmila & Tanya Ivanova-Sullivan. 2008. Lost in between: The case of Russian heritage speakers. *Heritage Language Journal* 6(1). 72–104.

Ivanova-Sullivan, Tanya. 2014. *Theoretical and experimental aspects of syntax-discourse interface in heritage grammars*. Leiden: Brill.

Kim, Ji-Hye. 2007. *Binding interpretations in Korean heritage speakers and L2 learners*. Urbana, IL: University of Illinois, Urbana-Champaign. (Doctoral dissertation).

Marrero, Victoria, Carmen Aguirre & María José Albalá. 2007. The acquisition of diminutives in Spanish: A useful device. In Ineta Savickiene & Wolfgang U. Dressler (eds.), *The acquisition of diminutives: A crosslinguistic perspective*, 155–181. Amsterdam, The Netherlands: John Benjamins Publishing.

Montrul, Silvina. 2008. Second language acquisition welcomes the heritage language learner: Opportunities of a new field. *Second Language Research* 24(4). 487–506.

Montrul, Silvina. 2010. Dominant language transfer in Spanish L2 learners and heritage speakers. *Second Language Research* 26(3). 293–327.

Nash, John C. 2014. On best practice optimization methods in R. *Journal of Statistical Software* 60(2). 1–14.

Nash, John C. & Ravi Varadhan. 2011. Unifying optimization algorithms to aid software system users: Optimx for R. *Journal of Statistical Software* 43(9). 1–14.

Nicoladis, Elena. 2002. The cues that children use in acquiring adjectival phrases and compound nouns: Evidence from bilingual children. *Brain and Language* 81(1–3). 635–648.

Nicoladis, Elena. 2006. Cross-linguistic transfer in adjective-noun strings by preschool bilingual children. *Bilingualism: Language and Cognition* 9(1). 15–32.

Paradis, Johanne, Elena Nicoladis & Fred Genesee. 2000. Early Emergence of structural constraints on code-mixing: Evidence from French–English bilingual children. *Bilingualism: Language and Cognition* 3(3). 245–261.

Pires, Acrisio & Jason Rothman. 2009. Disentangling sources of incomplete acquisition: An explanation for competence divergence across heritage grammars. *International Journal of Bilingualism* 13(2). 211–238. DOI: 10.1177/1367006909339806.

Polinsky, Maria. 2006. Incomplete acquisition: American Russian. *Journal of Slavic Linguistics* 14(2). 191–262.

Polinsky, Maria. 2008. Heritage language narratives. In Donna M. Brinton, Olga Kagan & Susan Bauckus (eds.), *Heritage language education: A new field emerging*, 149–164. London: Routledge.

Polinsky, Maria. 2018. *Heritage languages and their speakers*. Cambridge, UK: Cambridge University Press. DOI: 10.1017/9781107252349.

Qualtrics. 2021. Qualtrics research suite. Provo, Utah, USA. http://www.qualtrics.com.

R Core Team. 2021. R: A language and environment for statistical computing. Version 4.1.2. http://www.r-project.org.

Rakhilina, Ekaterina & Anastasia Marushkina. 2014. Non-calquing strategies in heritage Russian. Paper presented at the Second International Conference on Heritage/Community Languages, University of California, Los Angeles, CA.

Ravid, Dorit. 1998. Diminutive -i in early child Hebrew: An initial analysis. In Steven Gillis (ed.), *Studies in the acquisition of number and diminutive marking*, 149–174. Antwerp, Belgium: Antwerp University Press.

Rothman, Jason. 2009. Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International Journal of Bilingualism* 13(2). 155–163.

Rothman, Jason & Jeanine Treffers-Daller. 2014. A prolegomenon to the construct of the native speaker: Heritage speaker bilinguals are natives too! *Applied Linguistics* 35(1). 93–98.

Savickiene, Ineta. 1998. The acquisition of diminutives in Lithuanian. In Steven Gillis (ed.), *Studies in the acquisition of number and diminutive marking*, 115–135. Antwerp, Belgium: Antwerp University Press.

Scontras, Gregory, Maria Polinsky & Zuzanna Fuchs. 2015. Heritage language and linguistic theory. *Frontiers in Psychology* 6. Article 1545.

Scontras, Gregory, Maria Polinsky, C.-Y. Edwin Tsai & Kenneth Mai. 2017. Cross-linguistic scope ambiguity: When two systems meet. *Glossa* 2(1). 1–28.

Siegel, Jeff, Benedikt Szmrecsanyi & Bernd Kortmann. 2014. Measuring analyticity and syntheticity in Creoles. *Journal of Pidgin and Creole Languages* 29(1). 49–85.

Sorace, Antonella. 2000. Differential effects of attrition in the L1 syntax of near-native L2 speakers. In S. Catherine Howell, Sarah A. Fish & Thea Keith-Lucas (eds.), *Proceedings of the 24th Annual Boston University Conference on Language Development*, 719–725. Somerville, MA: Cascadilla Press.

Sorace, Antonella. 2011. Pinning down the concept of "interface" in bilinguals. *Linguistic Approaches to Bilingualism* 1(1). 1–33. DOI: 10.1075/lab.1.1.01sor.

Sorace, Antonella & Francesca Filiaci. 2006. Anaphora resolution in near-native speakers of Italian. *Second Language Research* 22(3). 339–368.

Tsimpli, Ianthi, Antonella Sorace, Caroline Heycock & Francesca Filiaci. 2004. First language attrition and syntactic subjects: A study of Greek and Italian near-native speakers of English. *International Journal of Bilingualism* 8(3). 257–277.

Wiese, Heike, Artemis Alexiadou, Shanley Allen, Oliver Bunk, Natalia Gagarina, Kateryna Iefremenko, Maria Martynova, Tatiana Pashkova, Vicky Rizou, Christoph Schroeder, Anna Shadrova, Luka Szucsich, Rosemarie Tracy, Wintai Tsehaye, Sabine Zerbian & Yulia Zuban. 2022. Heritage speakers as part of the native language continuum. *Frontiers in Psychology* 12. 717973.

Yip, Virginia & Stephen Matthews. 2000. Syntactic transfer in a Cantonese-English bilingual child. *Bilingualism: Language and Cognition* 3(3). 193–208.

Chapter 3

How different types of complexity can account for difficult structures in bilingual and monolingual language acquisition

Esther Rinke[a], Cristina Flores[b] & Jacopo Torregrossa[a]
[a]Goethe University Frankfurt [b]University of Minho

Certain linguistic structures are more challenging than others for bilingual speakers. This is true across different languages and language combinations. In this paper, we propose an account in terms of different types of linguistic complexity. Our argumentation derives from the results of a study based on a cloze test including 40 different linguistic structures of European Portuguese (EP). 180 children participated, all of them acquiring EP as a heritage language in Switzerland with different environmental languages (60 French-EP, 60 (Swiss) German-EP, and 60 Italian-EP bilinguals). The results show that the structures with the lowest accuracy rates are the same across the three groups. We single out four of these structures, namely, (i) *que* as a subject relative pronoun and as a consecutive conjunction, (ii) third person clitic pronouns in different forms and syntactic constellations, (iii) simple and contracted forms of prepositions, and (iv) the inflected infinitive in a concessive construction. We show that the difficulty of these structures reflects different forms of linguistic complexity: derivational complexity, memory-based learning, context dependency of rules and multiple form-function mappings. These forms of complexity cause difficulties also in monolingual acquisition.

1 Introduction

In this paper, we address the question whether and in which way the difficulties that heritage speakers (HSs) show with certain linguistic structures can be related to different types of linguistic complexity.

Esther Rinke, Cristina Flores & Jacopo Torregrossa. 2024. How different types of complexity can account for difficult structures in bilingual and monolingual language acquisition. In Maria Polinsky & Michael T. Putnam (eds.), *Formal approaches to complexity in heritage languages*, 43–71. Berlin: Language Science Press. DOI: ??

As a starting point, we provide the results of a study based on a cloze test focussing on a number of different structures of European Portuguese (EP) (Torregrossa et al. 2023). The test was completed by 180 bilingual children in the age span between 8 to 16 years with EP as their heritage language (HL) and different environmental languages (French, German, Italian).[1] The results show that certain structures are particularly difficult for the bilingual children, whereas others are unproblematic.

Because we find a very similar *hierarchy of difficulty* across the different language combination groups, we assume that the difficulties encountered by the child HSs are, in general, independent of the environmental language. The results of the abovementioned study challenge previous accounts which assign great importance to cross-linguistic influence as a factor determining deviances in bilingual production (see van Dijk et al. 2022 for a recent meta-analysis on cross-linguistic influences in bilingual morphosyntactic acquisition of diverse language pairs).[2]

Although we know that individual children's general proficiency is dependent on age and the amount of input that they receive in their HL (in terms of "quantity of language exposure"), it is still an open question why certain structures are more difficult to stabilize than others among bilingual as well as monolingual children.

We argue that the complexity of the target syntactic structures is crucially involved in defining the above-mentioned hierarchy of difficulty. However, it is very difficult to define what linguistic complexity actually means, because different notions and understandings of complexity exist in the literature. In order to approach our hypothesis, we will consider the four structures that caused the most difficulties for the children tested in Torregrossa et al.'s (2023) study, when

[1]Throughout the paper, we use the terms (simultaneous or early) bilingual children and heritage speakers to refer to the participants in our study. By "simultaneous bilinguals", we refer to the acquisition type, by "heritage speakers" we refer to the socio-political context of acquisition. HLs are minority languages spoken within families with a migration background. HLs are, thus, acquired in a bilingual context where another language is the official language of the society (majority/societal/environment language). Normally, as consequence of the acquisition setting, the majority language becomes the HSs' dominant language, but this is not always the case, i.e. language (im)balance is not taken as criterion to classify HSs.

[2]Note that this is not to say that CLI does not play any role in bilingual language acquisition. The argument goes the other way around: if we show that the same structures are complex and difficult for monolinguals and bilinguals with different language combinations, it becomes rather unlikely that CLI is the (one and only) relevant factor determining the difficulties in the acquisition of these structures by bilinguals. In any case, if CLI is argued to be a determining factor in HL development, this has to be unequivocally shown. It does not suffice to point to typological differences between the two languages of a bilingual speaker.

completing the cloze test in their HL. In particular, we will focus on i) *que* as a relative pronoun and consecutive conjunction, ii) clitic pronouns in different forms and syntactic constellations; iii) simple and contracted forms of prepositions and iv) the inflected infinitive in concessive sentences.

In order to show that the difficulty for the bilingual children indeed lies in the complexity of the structures (and is not related to bilingualism per se or cross-linguistic influence), we will first demonstrate that the respective structures that are difficult for bilingual children are also difficult for monolingual ones. In particular, we assume that lateness of a linguistic phenomenon in monolingual acquisition indicates its complexity for the learning/acquisition process. Based on previous proposals about complexity in monolingual language acquisition, we argue that complexity is a multifaceted notion. Our data allow us to identify the following types of complexity:

i. derivational complexity (layers of embedding, number of movement operations, instances of merge, e.g., in relative clauses)

ii. irregular and lexical forms that are memory-based (and not rule-based, e.g., lexically determined selection of "verb+preposition")

iii. context dependent rules (integration of syntactic and discourse knowledge, allomorphy dependent on phonological context, e.g., clitic allomorphy depending on the phonological context, contracted forms of prepositions in combination with definite articles)

iv. multiple form-function mappings (e.g., different functions of *que, por,* the use of the inflected infinitive in certain concessive clauses)

2 Empirical data: A hierarchy of difficulty in heritage EP

It is a well known fact that certain linguistic structures cause more difficulties for bilingual speakers than others, particularly in their non-dominant language (which is often, though not always, the HL). For example, bilinguals may show more problems than monolingual speakers with phenomena like gender assignment and agreement (Montrul et al. 2008), case marking (Polinsky 2006, 2008), pronoun realization and omission (Torregrossa et al. 2019, 2021), clitic allomorphs (Rinke & Flores 2014), subjunctive (Flores et al. 2017), and article realization (Montrul & Ionin 2010), just to mention a few.

In order to develop a proficiency assessment instrument for EP as HL, we constructed a cloze test, presented in detail in Torregrossa et al. (2023). In general,

cloze-tests are considered to be integrative assessment tools, because the participants have to access their linguistic knowledge to reconstruct the missing gap in the test (Chung & Ahn 2019).

The study was conducted in Switzerland, with bilingual children with different language combinations (Portuguese-French, Portuguese-German and Portuguese-Italian), as depending on the Swiss canton of residence.

2.1 Participants

The study included 180 child HSs, 60 children for each language combination. Most of the children were born in Switzerland or emigrated there early in life. All participants acquired Portuguese from birth and the environmental language as a second first or early second language. Their age ranged from 8;6 to 16 years (M: 11;7; SD: 1;10). The study was conducted in cooperation with the *Camões* Institute, where all participants attended HL classes weekly. The cloze test was conducted as an untimed written task during a HL class.

Switzerland is an ideal place to conduct this type of study, because there lives a fairly large community of Portuguese-speaking migrant families. Their children acquire the heritage language, EP, in the context of three different dominant environmental languages: French/German/Italian. In addition, Switzerland has a tight network of Portuguese HL classes, offered by the Portuguese Institute for the maintenance and development of Portuguese abroad (*Instituto Camões,* see de Lourdes Gonçalves & Vinzentin 2021).

In addition to the data presented in Torregrossa et al. (2023), we collected data from 23 monolingual Portuguese children in the ages of 12-13 years (M: 12;3; SD: 0;5) for the sake of the present discussion. They completed an online version of the same cloze test.

2.2 Test and coding methodology

The cloze-test is based on a short narrative modelled after the B3 story of the Edmonton Narrative Norms Instrument (ENNI; Bongartz & Torregrossa 2020, Schneider et al. 2005). The test includes 40 gaps with a variety of structures tapping into different linguistic domains: nominal morphology, verbal morphology, (contracted and non-contracted) prepositions, different types of complementizers, (clitic) pronouns in different syntactic constellations, definite and indefinite articles, and lexical knowledge. For functional words, we deleted the whole word or provided the initial letter in order to facilitate completion and restrict the number of possible answers. For content words, we provided the first half of the word

(as is usually done in c-tests) for the same reasons. The results were coded according to the following four options: correct, incorrect, missing, or not expected but correct. For the analysis, we considered the correct and unexpected (but correct) answers as "correct" (1) and the incorrect and missing answers as "incorrect" (0).

Needless to say, different structures can be difficult for different reasons. In order to be able to differentiate between structure-related factors and other causes, we also collected information on the language background of the children (age of onset to the second language, quantity of input, length of attendance of HL classes, etc.). Concerning the relevance of these factors we refer the reader to Torregrossa et al. (2023).

2.3 Results

In Table 1 we report the overall results for monolingual and bilingual children. Since the data collection method is different (online vs. in paper form) and the monolingual children's age range is more limited, the results have to be interpreted with caution. Nonetheless, they provide us with additional evidence related to a hierarchy of difficult structures, which we argue to hold for all children, independently of their being monolingual or bilingual.

Table 1: Accuracy rates of monolingual and bilingual children

	Bilinguals	Monolinguals
Overall accuracy rate	4635/7200 (64.4%)	843/920 (91.6%)
Max	170/180 (94.4%)	23/23 (100.0%)
Min	51/180 (28.3%)	9/23 (39.1%)

Across all 40 target structures, the 180 bilingual children show an accuracy rate of 64.4% (4635/7200; max. 170/180 (94.4%)/min. 51/180 (28.3%)). The accuracy rate of the monolinguals is 91.6% (843/920; max 23/23 (100%)/min. 9/23 (39.1%)). A closer look at the results reveals that some of the structures are indeed particularly challenging for the children. The following structures received the lowest accuracy rates:

2.3.1 *que* as a relative pronoun and consecutive conjunction

The element *que* has a number of different functions in EP and occurs in different types of subordinating constructions. It may serve as a complementizer introduc-

ing a complement clause (1a), a relative pronoun (1b) or a consecutive adverbial conjunction (1c).

(1) a. [item 18]
Ele pensa **que** pode ir buscar um balão para a sua amiga.
he thinks that can go bring a balloon for the his friend
'He thinks that he can bring a balloon for his friend.'

b. [item 12]
Mas sem querer, o coelhinho larga o balão, **que** voa para longe.
but without wanting the rabbit releases the balloon that flies to far away
'Without wanting it, the rabbit releases the balloon that flies away.'

c. [item 14]
A cadelinha está tão zangada **que** começa a gritar (...).
the little dog is so angry that starts to shout
'The little dog is so angry that she starts to shout (...).'

Table 2 shows that the constructions mentioned in (1b) and (1c) received low accuracy rates in the bilinguals' cloze test, which indicates that they are difficult for the children.

Table 2: Accuracy rates of constructions with different types of *que*-subordinators

Example with gap	expected item	grammatical category	accuracy (bilingual's)
Mas sem querer, o coelhinho larga o balão, ___ voa para longe. (see 1b)	[que]	subject relative pronoun	28.3% (51/180)
A cadelinha está tão zangada ___ começa a gritar e a discutir em voz alta com o seu amigo. (see 1c)	[que]	adverbial consecutive conjunction	49.4% (89/180)

2.3.2 Third person clitic pronouns in different forms and syntactic constellations

Clitic pronouns in EP are marked for a number of different morphological features (e.g. gender/number/case) and can occur as simple clitics (2a) or contracted forms (clitic allomorphs) (2b–c). Clitic allomorphs are allomorphic forms of the clitic pronouns *-o(s)*/*-a(s)* that change for phonetic reasons due to the ending of the verb form which they attach to. For instance, in examples (2b) and (2c), the clitic *-o* (singular, *-os* plural) changes its form to *-lo(s)* because the verb form ends with the consonant /r/.

(2) a. [item 15]
Assustado, este ouve-**a** *a gritar.*
scared this-one hears-her to cry
'Scared, he hears her crying.'

b. [item 7]
O coelhinho quer tirá-**lo**.
the little rabbit wants take-it
'The little rabbit wants to take it.'

c. [item 33]
e pergunta-lhe, se poderia ajudá-**los.**
and asks-him if could help-them
'And he asks him whether he could help them.'

Table 3 (p. 50) shows the accuracy rates associated with the mentioned structures.

2.3.3 Simple and contracted forms of prepositions

Many Portuguese prepositions can occur in a contracted form with a definite determiner. Examples are the prepositions *em* + *a/o* (in + the fem./the masc.) = *na*/*no* and *por* + *a/o* (for/through+the fem./the masc.) = *pela*/*pelo* (3a). Besides prepositional phrases with an adverbial contribution, the preposition *por* marks the agent of a passive verb in EP, as shown in (3b). Prepositions in combination with verbs can also lead to a new verb meaning, which is semantically opaque, in the sense that it does not derive compositionally from the meaning of the verb and the one of the preposition. In example (3c), the combination of the verbs *ir* and *ter* (go + have) with the preposition *com* (with) leads to the interpretation 'go to see'.

Table 3: Accuracy rates of (clitic) pronouns in different forms and syntactic constellations

Example with gap	expected item	grammatical category	accuracy (bilingual's)
Assustado, este ouve-_ a gritar. (see 2a)	[a]	clitic pronoun (feminine, singular, accusative)	40% (72/180)
O coelhinho quer tirá-__ (see 2b)	[lo]	clitic pronoun (allomorph, masculine, singular, accusative)	38.8% (70/180)
e pergunta-lhe, se poderia ajudá-__ (see 2c)	[los]	clitic pronoun (allomorph, masculine, plural, accusative)	44.4% (80/180)

Table 4: Accuracy rates of (simple and contracted forms of) prepositions

Example with gap	expected item	grammatical category	accuracy
decidem ir passear p__ floresta (see 3a)	[pela]	preposition (contraction: *por + a)*	40% (72/180)
A mãe ouve com atenção o relato feito __ ele, (see 3b)	[por]	preposition (passive agent)	42.7% (77/180)
vai ter __ o coelho vendedor, e pergunta-lhe pelo preço do balão. (see 3c)	[com]	preposition (in fixed verbal expression)	53.3% (96/180)

(3) a. [item 2]
decidem ir passear **pela** *floresta*
decide go walking through-the forest
'They decide to go for a walk through the forest.'

b. [item 34]
A mãe ouve com atenção o relato feito **por** ele,
the mother listens with attention the report made by him
'The mother listens with attention to his report.'

c. [item 35]
vai ter **com** o coelho vendedor, e pergunta-lhe pelo preço
goes have with the rabbit salesman and asks-him for-the price
do balão.
of-the balloon
'He goes to see the salesman rabbit and asks him for the price of the balloon.'

Table 4 shows that prepositions are difficult for the bilingual children, in particular in contexts like (3a) and (3b).

2.3.4 Inflected infinitives in concessive constructions

EP possesses a special syntactic construction: the inflected infinitive. The construction is relatively frequent, especially in final clauses introduced by the preposition *para* as in (4a). The inflected infinitive occurs also in concessive clauses introduced by *apesar de* ('although', as in 4b).

(4) a. Os pais foram à livraria para **comprarem** os livros
the parents went to+the book store to buy+3PPl the
escolares novos.
school books new
'The parents went to the book store to buy the new school books.'

b. [item 28]
Apesar de eles **pedirem** com muita educação, ...
despite of they ask with much education, ...
'Although they asked delicately, ...'

Table 5 shows that the inflected infinitive in concessive constructions also belongs to the difficult structures, with less than 50% accuracy.

Table 5: Accuracy rates of the inflected infinitive in concessive constructions

Example with gap	expected item	grammatical category	accuracy
Apesar de eles pedir__ com muita educação, (see 4b)	[pedirem]	inflected infinitive 3P Plural	47.2% (85/180)

Taking into account the results per language combination, we find that the abovementioned structures are associated with low accuracy rates across the three groups considered in this paper, as shown in Table 6.

Table 7 reports for each language combination group, the 12 structures with the lowest accuracy rates in the cloze-test. We highlighted in bold the structures that were common across the three language combination groups. Notably, 11 out of the 12 structures were the same for the three groups. The Portuguese-German and Portuguese-Italian children share all 12 structures, even if in a slightly different order of accuracy. The list of structures related to the Portuguese-French children included the irregular plural noun phrase *balões* ('balloons'), instead of the preposition *com* ('with').

Table 7 shows that the structures that are most difficult for the bilingual children in this study are very similar across the three language combination groups. The fact that the nine linguistic structures discussed here (see i–iv above) belong to the 12 most difficult structures independently from the contact language suggests that the bilingual children's difficulties with these structures are unlikely due to cross-linguistic influence. If it is true that these structures are associated with a complex learning task, they should be difficult for monolingual children, too.

As mentioned above, we collected data from 23 monolingually raised children, living in Portugal. The data are not fully comparable, because the cloze test was conducted online (during the COVID-19 restriction period) and included only 12-13 years-old children. As expected, the rate of accuracy was much higher in this group. Nonetheless, even in this case, we identified some difficult structures that did not reach ceiling performance. Table 8 shows the three structures with the lowest accuracy rates, which overlap with the structures listed in Table 7.

These data suggest that the structures that are most difficult for bilingual children are also challenging in monolingual acquisition. However, since the data

Table 6: Accuracy rates for the most difficult structures across the language combination groups

expected item	grammatical category	Ptg./French mean: 61.5% (1476/2400)	Ptg./German mean: 66.1% (1588/2400)	Ptg./Italian mean: 65.4% (1571/2400)
[que]	subject relative pronoun	15% (9/60)	33.3% (20/60)	36.6% (22/60)
[que]	adverbial consecutive conjunction	38.3% (23/60)	60% (36/60)	50% (30/60)
[lo]	clitic pronoun (allomorph, masculine, singular, accusative)	33.3% (20/60)	43.3% (26/60)	40% (24/60)
[a]	clitic pronoun (feminine, singular, accusative)	38.3% (23/60)	46.6% (28/60)	35% (21/60)
[los]	clitic pronoun (allomorph, masculine, plural, accusative)	33.3% (20/60)	51.6% (31/60)	48.3% (29/60)
[pela]	preposition (contraction: *por + a)*	42.6% (25/60)	43.3% (26/60)	35% (21/60)
[por]	preposition (passive agent)	38.3% (23/60)	43.3% (26/60)	46.6% (28/60)
[com]	preposition (in fixed verbal expressions)	55% (33/60)	53.3% (32/60)	50% (30/60)
[pedirem]	inflected infinitive, 3P plural	38.3% (23/60)	55% (33/60)	46.6% (28/60)

Table 7: List of the 12 most difficult structures and the associated accuracy scores for each language combination group.

	Ptg./French		Ptg./German[a]		Ptg./Italian	
1	[que]$_{\text{SUBJ\_REL}}$	(15%)	[que]$_{\text{SUBJ\_REL}}$	(33.33%)	[repara]$_{\text{PRES.3.SG.}}$	(31.67%)
2	[lo]	(33.33%)	[pela]	(43.33%)	[pela]	(35%)
3	[los]	(33.33%)	[lo]	(43.33%)	[a]	(35%)
4	[que]$_{\text{\_CONS}}$	(38.33%)	[por]	(43.33%)	[que]$_{\text{SUBJ\_REL}}$	(36.67%)
5	[a]	(38.33%)	[a]	(46.67%)	[lo]	(40%)
6	[pedirem]	(38.33%)	[lhe]	(50%)	[lhe]	(41.67%)
7	[por]	(38.33%)	[vem]	(51.67%)	[vem]	(46.67%)
8	[balões]	(38.33%)	[los]	(51.67%)	[por]	(46.67%)
9	[pela]	(41.67%)	[repara]$_{\text{PRES.3.SG.}}$	(53.33%)	[pedirem]	(48.33%)
10	[lhe]	(41.67%)	[pedirem]	(55%)	[los]	(48.33%)
11	[vem]	(55%)	[com]	(55%)	[que]$_{\text{\_CONS}}$	(50%)
12	[repara]$_{\text{PRES.3.SG.}}$	(60%)	[que]$_{\text{\_CONS}}$	(60%)	[com]	(50%)

[a]At first sight, the Portuguese/German group seems to show higher accuracy rates. However, the statistical analysis in Torregrossa et al. (2023) clearly shows that the language combination did not have an effect on response accuracy.

Table 8: Three most difficult structures for 12/13-years old monolinguals

	Portuguese monolinguals	
1	[que]$_{\text{SUBJ\_REL}}$	(39.1%)
2	[pela]	(56.5%)
3	[a]	(78.3%)

collection method is different and the children's age range is much more limited, we will support this hypothesis by relying also on existing studies on the L1 acquisition of the phenomena under discussion in the next section.

3 A look at monolingual acquisition

In this section, we look briefly at the main findings reported in previous literature on the L1 acquisition, in Portuguese (and in other languages), of the target structures, namely at the acquisition of conjunctions, clitic pronouns, prepositions, and concessive connectors requiring the inflected infinitive.

3.1 Subordinators and subordinate clauses

Amid the various complementizers introducing subordinate clauses, *que* seems to be one of the first to appear in child EP, in complement clauses (see 1a), following the emergence of complement clauses with infinitives (Santos 2017). However, as already mentioned in §2, *que* introduces different types of subordinate clauses and we know that not all of them are acquired at the same time in EP, as in many other languages (Soares 1998).

Research on the acquisition of EP, in particular the study conducted by Soares (1998), has shown that relative clauses are amongst the latest types of subordinate clauses to appear in child speech (see also Vasconcelos 1995). This has been shown also for other languages. For instance, Bloom et al. (1980) and Dromi & Berman (1986) proposed that, in English and Hebrew complement clauses emerge first, followed by adverbial clauses, and lastly, relative clauses (but see Penner 1995 for a different order in Swiss German). Various explanations have been proposed to account for the order of acquisition of different subordinate clauses (Bowerman 1979). Traditionally, it is attributed to different degrees of embedding: The structure that has fewer layers of embedding is less complex and, therefore, easier to acquire. This would be the case of complement clauses, which are selected by the matrix verb in the same fashion as any other verbal complement. Adverbial clauses are not selected directly by the verb, but they involve one layer of embedding. Thus, they emerge later than complement clauses, but earlier than relative clauses, which involve both embedding and movement. Since relative clauses are the most complex structures in terms of embedding, they would be the last structure to emerge. In fact, EP children have difficulties in producing and comprehending relative clauses until school-age (Vasconcelos 1995).

Armon-Lotem (2005) argues that it is necessary to look not only at the timing of emergence of certain structures, but at the timing of its complete stabilization, since a structure is only completely stabilized in the child's grammar when all the associated features are acquired and the structure is used in all relevant contexts. This explains cross-language differences and further distinctions within each type of subordination considered above. For instance, in EP there are different timings of acquisition of complement clauses due to different timings of acquisition of verbal semantics and verbal mood (Jesus et al. 2019). For relative clauses, it has been shown that right-embedded clauses emerge earlier than middle-embedded ones (Vasconcelos 1995) and that subject relative clauses are easier to acquire and process than object relative clauses (Costa et al. 2011).

Furthermore, a typical property associated with the acquisition of subordination is the omission of the complementizers, which starts at a pre-conjunctional

period, but is prolonged throughout the acquisition process until later stages of acquisition (Armon-Lotem 2005, Soares 1998).

For the purpose of our discussion, the main observation to retain is that, in child EP, complement clauses stabilize earlier than adverbial clauses and these stabilize earlier than relative clauses. A frequent feature of child subordination is the omission of the complementizer.

3.2 Clitic pronouns in different forms and syntactic constellations

It is a well-established fact that EP has a rich pronominal system. In addition to allowing for the use of strong and clitic pronouns, EP is also a null object language. This means that children acquiring EP have to acquire the conditions of use of strong pronouns, clitics or clitic omission, including null objects, VP ellipsis or other types of object omissions. Several studies focusing on the production and comprehension of clitics and null objects by monolingual EP children demonstrate that they go through a prolonged stage of object omission and stabilize knowledge of the pronominal system very late (at school-age; see Costa & Lobo 2007, 2009, Costa et al. 2009, 2012, Flores et al. 2020, among others). It is argued that the overuse of null objects is caused by children's difficulties in assigning the correct interpretation to different types of object omissions available in the target grammar (pro, variable, VP-ellipsis, null object; cf. Costa et al. 2012). Due to the complexity of the pronominal system, EP L1 children omit objects to a higher degree and for a longer period of time than children acquiring other Romance languages that have clitics, or even other null object languages (Varlokosta et al. 2016). Despite this delay, EP children show early pragmatic knowledge of pronoun use (Costa et al. 2009, Flores et al. 2020). This indicates that the prolonged non-adult-like interpretation and production of pronouns lies, on the one hand, in the acquisition of the feature composition of the null objects and, on the other hand, in the acquisition of some syntactic and morphological features of clitics.

A syntactic property of clitics that has been shown to stabilize late in L1 EP is clitic placement. Differently from other clitic languages (and even differently from Brazilian Portuguese), several syntactic constraints determine the pre- or postverbal position of the clitic pronoun in EP. In particular, the preverbal position (proclisis) is stabilized very late in L1 acquisition (by age 7, see Costa et al. 2015).

In addition to the late acquisition of the properties constraining the realization vs. omission of the object pronoun and its placement, certain morphological features also stabilize only at school-age. A case in point is mesoclisis, i.e. the occurrence of allomorphic clitic forms in the middle of the verb form (*e.g.* 1P Sing.

future form *eu vê-**lo**-ei* 'I will see it/him') and clitic allomorphy in postverbal (enclitic) position (*tirá-**lo*** 'take it', *ajudá-**los*** 'help them'; see 2b and 2c). The target-like use of these structures is sensitive to formal instruction and shows variation in colloquial Portuguese (see Catalão 2011, Santos 2002 and Batalha 2018 for an analysis of Portuguese school-aged children's knowledge of clitic pronouns).

3.3 Prepositions

Prepositions are a heterogeneous category that includes elements with lexical meaning (e.g. spatial prepositions) and semantically vacuous elements functioning as grammatical markers (e.g. the dative preposition *a*). Lexical prepositions have their own lexical entry, whereas non-lexical prepositions have undergone some form of grammaticalization and have a purely syntactic function or they occur in fixed phrases (Rauh 1993, van Riemsdijk 1990). This split into lexical vs. functional prepositions (or a continuum from more lexical to more functional prepositions) is mirrored in the process of acquisition of languages with a prepositional system. For example, Littlefield (2009) argues that in L1 English, lexical prepositions emerge early and show a steady, relatively rapid increase in child speech over time. Inversely, pure functional prepositions (*e.g.,* 'of') emerge later and their production is limited and often not target-like in the first stages of acquisition. The same seems to hold for Portuguese, even though research on the acquisition of prepositions in Portuguese is scarce (Teodoro 2020).

A further characteristic of prepositions which is visible across several languages is the contraction of the preposition with other elements, such as pronouns or articles. In Portuguese, the contraction of the preposition with the definite article (see §2.3) is almost categorical, with only few syntactic contexts representing an exception. In addition to always requiring gender and number marking, there are contractions that change the stem (*e.g. por + a = pela* 'through-the') and contractions that involve only the deletion of the final vowel (e.g. *de + a = da* 'of-the'). Due to the absence of research on the acquisition of prepositional contractions in L1 acquisition of Portuguese, we will resort to studies on L2/L3 research (Brito 2018, Picoral & Carvalho 2020). In a study with Spanish and English native/heritage speakers learning Portuguese as L3, Picoral & Carvalho (2020) show that speakers are more likely to realize contractions with the preposition *a* and that the contraction of the preposition '*por* + definite article' is the most difficult to acquire. Furthermore, the acquisition path seems to be independent of the speakers' L1.

As for the preposition *por*, in addition to a spatial meaning, it has also the pure grammatical function of introducing the agent in passive sentences (as by-

phrase), either in contracted form or not, depending on the presence or absence of a definite article, respectively. It has been argued that *by*-phrases of passive sentences are generally problematic for children (Fox & Grodzinsky 1998). This difficulty may be due to several factors, including the type of passive sentence (e.g., long or short-actional passives; see Armon-Lotem et al. 2016), the agentivity of the predicate (Estrela 2015) and the above-mentioned difficulty for children to use semantically vacuous prepositions.

3.4 Inflected infinitives in concessive constructions

We know from studies on spontaneous child speech that inflected infinitives emerge early in EP (Santos 2017), i.e., by the age of two years. However, at an initial phase, they only occur in final clauses introduced by *para* (Santos et al. 2013). Only later (i.e., by the age of three years), they occur in complements of perception verbs (Santos et al. 2016). This means that, even though the inflected infinitive is available to EP children from early on, the different contexts that allow its use are acquired gradually, which depends on both syntactic and lexical constraints. In fact, some contexts requiring the use of an inflected infinitive are acquired very late, i.e. in school-age. This is the case of the concessive structure *apesar de* ('although').

According to Costa (2006), the concessive connector *apesar de* is stabilized very late in EP (i.e. only by the age of ten years, similar to the stabilization of *although* or *whereas* in English, see Diessel 2004). Costa (2006) argues that this late acquisition is caused by three different, but interacting factors. The first factor is frequency: The connectors *apesar de* and *embora* are produced significantly less by adults than the adversative connector *mas*. However, frequency per se does not explain the late acquisition of this structure. The late stabilization of concessive connectors may be related to the fact that they occur only in subordinate clauses and most of them require the use of the subjunctive, which is also stabilized late in EP.

4 The role of linguistic complexity

The discussion in §3 has shown that the different structures under consideration are not only difficult for bilingual children, but are also mastered relatively late by monolinguals. If these structures take time to be acquired in monolingual language acquisition, we expect to find an effect of age in the bilingual group as well. Thus, we ran a statistical analysis to assess the effect of the bilingual children's age on the acquisition of the most difficult structures. We considered the

nine structures which are relevant for the present paper (see i–iv in §2). As we mentioned in §1, the age range of the participants is relatively large (i.e., from 8;6 to 16 years; *M*: 11;7; SD: 1;10). We ran a binomial logistic regression with accuracy as dependent variable (0 = inaccurate, 1 = accurate) and age as fixed effect. The model showed a significant effect of age ($\beta = 0.64$, SE $= 0.06$, $z = 10.30$, $p < 0.001$). This shows that bilingual children's knowledge of difficult structures improves with age. In this sense, bilinguals behave just like monolinguals, even if they may need more time to acquire difficult structures. In this sense, it is possible that the structures that are not mastered by younger bilingual children are exactly the structures that emerge late in monolingual language acquisition. In other words, these structures are 'complex' for bilinguals and monolinguals alike, as shown by their late timing of acquisition across the board. Since it is often observed that bilinguals show a more protracted development, i.e., they acquire some structures in later age spans than monolinguals, we assume that bilinguals just need some more time to catch up with their monolingual peers (see Schulz & Grimm 2019, Tsimpli 2014 for similar considerations). In the remainder of this paper, we intend to discuss why certain structures are associated with a more complex learning task than others.

4.1 Notions of linguistic complexity

In the literature, complexity in acquisition has been explicitly defined and implicitly assumed in many different ways. From a syntactic perspective, it has been assumed that children initially prefer more syntactically economical structures over less economical ones; i.e. structures involving less layers of embedding over structures involving more layers of embedding, or structures involving less movement operations over structures involving more movement operations (Hamann 2006, Rizzi 1990, 2000). Jakubowicz (2003) proposes that computational complexity affects child language development, leading children to produce less complex structures in a target-like way earlier than more complex structures (see also Jakubowicz & Nash 2001). The author proposes the following *Derivational Complexity Metric*.

(5) Derivational Complexity Metric (DCM, Jakubowicz 2011)
- Merging α_i n times gives rise to a less complex derivation than merging α_i $(n + 1)$ times.
- Internal Merge of α gives rise to a less complex derivation than Internal Merge of $\alpha + \beta$.

For example, with respect to wh-questions, the DCM predicts, "that the child is sensitive to the number of times that a copy of the wh-element must be merged to satisfy a computational requirement and to the number of constituents that may (or must) undergo Internal Merge (here under: IM)" (Jakubowicz 2011: 340; see also Soares 2003 with respect to the acquisition of wh-questions in EP).

The notion of complexity presented so far is motivated syntactically. Another way of defining complexity is more morphologically oriented and based on the observation that children tend to overregularize morphological endings. Clahsen et al. (2002) argue for a dual-mechanism model between rule-based (regular) and memory-based irregular representations for morphologically complex words. In their study, children acquiring Spanish verb morphology overapply regular paradigms to verbs that require irregular forms but not vice versa. The authors argue that "...the onset of overregularizations is syntactically triggered, by the requirement to generate a fully specified finite verb form in every sentence, in conjunction with lexical gaps or retrieval failures for irregulars. Overregularizations gradually decrease over time when children get older and memory traces for irregulars are becoming stronger and the children's ability to retrieve them is becoming more reliable" (Clahsen et al. 2002: 618). Coming back to the issue of complexity in acquisition, these results suggest that regular syntactic or morphological rules are less complex than irregular forms, which have to be memorized and stored in the lexicon based on individual forms (and their frequency) in the input. Hence, the acquisition of rules that are applied regularly seems to be less costly than memory-based lexical learning.

The morphological rule mentioned in the previous example is based on a syntactic requirement (namely to generate a fully specified finite verb form) that applies independently of the context (i.e., the situation in which the sentence is uttered) and, in principle, concerns every sentence. However, this is not the case for each morphological or syntactic rule. We would like to add another type of complexity which lies in-between rule-based regular and memory-based irregular representations, namely cases in which a rule is applied depending on a specific (discourse or phonological) context. We suggest that this also involves a two-step learning/acquisition process: acquiring the rule and understanding in which context it applies and in which context it does not.

One example that has been mentioned in a number of studies is context dependency of a form which is related to previous discourse. In languages with null and overt pronouns, this concerns, for example, the decision whether a pronoun has to be overtly realized or can remain phonologically null. It has been suggested that in null subject languages, bilingual speakers tend to overrealize pronouns

compared to monolingual speakers and may fail to accurately differentiate between the two forms in interpretation tasks. Sorace et al. (2009: 464) argue that this is a result of the complexity of the task: bilinguals have more difficulties in integrating different sources of information. According to Sorace (2011), the differences between monolingual and bilingual populations relate to bilingualism per se and, in particular, to the allocation of general cognitive resources to bilingual processing. However, the complexity of integrating syntactic information and discourse information represents a complex learning task also for young monolinguals (as shown for Portuguese by Lobo & Silva 2016, Rinke & Flores 2018) and may result in a protracted development of such phenomena. For example, Tuller et al. (2011) observe that in French, 3rd person accusative clitics are difficult among young TD (=typically developing) children and AD (atypically developing) speakers after childhood. The authors claim that the

> complexity of object clitics is the result of a combination of several properties, the first of which is their non-canonical position. […] Summarizing, the production of accusative clitics includes the following properties: movement to a non-argument position, clustering with nominative clitics, and reference to a non-local antecedent. Production of a third person accusative clitic involves the following additional properties: establishing non-discourse-dependent reference, agreement in both number and gender, but not animacy, and, potentially, licensing of a null clitic (conditional on both lexical and discourse restrictions). They are thus complex (morpho)syntactically, in terms of movement (whichever analysis of clitic constructions is adopted) and agreement, and mastering their usage (knowing whether they can be null or not) requires adhering to lexical idiosyncrasies and discourse/pragmatic conditions. (Tuller et al. 2011: 427f.)

A similar observation applies to 3rd person object clitics in EP, whose production is associated with the same degree of complexity as ascribed by Tuller et al. (2011) to French clitics. In addition, EP allows for 3rd person null objects in similar syntactic and discourse contexts as clitics. Therefore, the acquisition of the target-like distribution of object clitics and null objects in a null object language like EP and, hence, of the discourse-appropriate production of clitics is more challenging than the acquisition of clitics in non-null object languages (Costa & Lobo 2007, Flores et al. 2020, Varlokosta et al. 2016).

In addition, EP clitics show allomorphy in certain phonological contexts, as described above. Allomorphic variation represents another form of linguistic complexity. It has been shown, for example, that allomorphic variation of English

past tense forms (e.g., "-t for verbs such as *chase*, -d for forms such as *crave* and /əd/ for verbs such as *recite*") slows down morphological development (O'Grady et al. 2010: 369). O'Grady et al. (2010) also mention homophony as a factor determining morphological development in first language acquisition.

> Whereas the word *the* functions only as a determiner in English, the suffix -s can be used to mark any one of three things: plural number in nouns, third person singular in verbs, or possession. The resulting complication in the relationship between form and meaning may impede acquisition.
>
> (O'Grady et al. 2010: 369)

We assume that in general, multiple form-function mappings (e.g. allomorphy, homophony) give rise to complexity in acquisition and may cause difficulties or a slow down in development. To conclude, we identified the following types of linguistic complexity in first language acquisition: i) derivational complexity (layers of embedding, number of movement operations, instances of merge); ii) irregular and lexical forms that are memory-based (and not rule-based); iii) context dependent rules (integration of syntactic and discourse knowledge or allomorphy depending on phonological context) and iv) multiple form-function mappings. In the next section, we will discuss how these notions of complexity apply to the "hierarchy of difficulty" discussed in §2.

4.2 Towards an explanation of the hierarchy of difficulty

In this section, we would like to come back to the phenomena mentioned in sections 2 and 3 that were the most challenging linguistic structures for the children and explore to what extent their difficulty can be related to the above mentioned notions of linguistic complexity.

As already discussed in §2, the item *que* as a relative pronoun and as a consecutive complementizer belonged to the constructions with the lowest accuracy rates across the different language combination groups. It is interesting to contrast these two structures with the declarative complementizer *que* illustrated in (1a), which is associated with a high accuracy rate of 70.5% (vs. 28.3% for the relative pronoun and 49.4% for the consecutive complementizer). As shown in §3, the different accuracy rates for the different types of *que* correspond to the order of acquisition of the different instantiations of *que* in monolingual EP: the declarative complementizer is acquired first in child EP, followed by *que* introducing adverbial clauses, followed in turn by relative clauses, some of which may also emerge at school age. Even the 12–13 years-old monolingual children

showed low rates of accuracy in association with the relative pronoun *que*. In §3, we mentioned that the difference between the various types of subordinate clauses (complement clauses selected by the verb, adjoined adverbial clauses and relative clauses) can be accounted for in terms of degrees of derivational complexity, involving, for example, embedding (in concessive clauses) or embedding and movement (e.g., in relative clauses). An additional factor contributing to the complexity of the structures at stake is the multiple form-function mapping of *que* in EP (one form with several functions), namely the homophony of *que* as a conjunction of complement and adverbial clauses, as an interrogative or a relative pronoun or an interrogative determiner.

Third person clitic pronouns represent another area of difficulty in the cloze tests among the bilingual children. As mentioned in §3.2, these structures are also very challenging for EP monolingual children and acquired successfully only at school age. Clitics are difficult for a number of different reasons. In addition to potential (syntactic) derivational complexity (if we assume a movement analysis for clitics), clitics are morphologically complex because they involve allomorphy in EP. Depending on the phonological context, the form of the clitic may change. For example, following the –r ending of infinitives, the clitic *o* (acc. masc. sing.) is realized as *–lo* (see example 2b, c); after the nasal *–m* (e.g., 3rd person plural finite verb forms), *o* surfaces as *–no*. As discussed in the previous section, such rules are complex for different reasons: they are context dependent (therefore involving a two step learning process) and there is no direct form-function mapping (because different forms have the same function and realize the same morphological features). A third factor contributing to the complexity of clitic pronouns is their discourse dependency, since the appropriate use of clitics (as well as null objects and full noun phrases) is dependent on discourse constraints (Flores et al. 2020).

The third phenomenon discussed in Sections 2 and 3 are prepositions in different shapes and constellations. We saw that in the cloze test, bilingual children, but also monolinguals, show low accuracy with contracted forms of prepositions. In addition, the bilingual children have also problems with the preposition *por* introducing passive agents and the lexically selected preposition *com*. First of all, contracted forms of prepositions are derived based on a context-dependent rule (only in combination with definite articles, not with indefinite ones or bare nouns). Assuming a Distributed Morphology approach, Ximenes (2004) states that contractions of prepositions are the results of a two-step morphological process: "two operations happening in the morphological component: merger followed by fusion." (Ximenes 2004: 182). As already discussed in §3, *por* as a preposition marking the agent of a passive sentence is generally problematic for younger

children and complex, because it is a functional and, hence, a semantically vacuous preposition. The homophony with the lexical preposition *por* marking a spatial meaning leads to multiple form-function mapping and may contribute to the complexity of this preposition as well. The complexity of the preposition *com* in combination with the verb *ter* (see example 3c) has a different source. In this context, the preposition contributes to the formation of a new verb – a process that is very productive in EP (e.g. *acabar de* 'finish (of)', *acabar com* 'destroy', *acabar por* 'end up by'). Crucially, the combination of a verb and a preposition is semantically opaque and can only be acquired through a memory-based lexical learning process.

The fourth phenomenon that was associated with some difficulty for the bilingual children in the cloze test was the inflected infinitive in combination with the concessive connector *apesar de*. As mentioned in §3, EP monolingual children do not exhibit any difficulty in the use of inflected infinitives. However, concessive connectors are acquired late and exhibit a similar degree of complexity as other conjunctions introducing adverbial phrases. When introducing a clause, *apesar de* occurs only in combination with inflected and uninflected infinitives. In more formal registers, we find the (more complex) construction *apesar de que,* which introduces finite subordinate clauses that require the indicative or the subjunctive mood, which is another property of EP which is acquired relatively late in L1 acquisition. In addition to these forms that belong to the standard register, we find also the occurrence of *apesar que* in association with the indicative mood in colloquial speech. Furthermore, *apesar de* may also introduce a NP with concessive meaning, instead of a clause (e.g., *Apesar da chuva, eles foram passear.* 'Despite the rain they went for a walk.'). Hence, the difficulty related to the use of the inflected infinitive in the cloze test does not depend on the structure itself, but results from its combination with the concessive connector *apesar de*, which is acquired late and can introduce different structures. In particular, its alternation with a finite verb in the same context, as in the use of the indicative with the non-standard *apesar que*, may increase the difficulty of the acquisition task. Actually, the most frequent error committed with this item was the replacement of the inflected infinitive with the finite 3rd person plural indicative form *pediram.*

5 Summary and conclusions

The present line of argumentation derives from the observation that some linguistic structures cause more difficulties for bilingual speakers than others, especially in their non-dominant HL. We aimed to show that the difficulty of certain structures is related to different types of linguistic complexity. A cloze test conducted

with 180 EP heritage children in Switzerland – divided into three groups with 3 different environmental languages (French, German and Italian) – revealed that the children exhibited particular difficulties with some structures, including relative pronouns and consecutive conjunctions, clitic pronouns in different forms and syntactic constellations, some simple and contracted forms of propositions and inflected infinitives in concessive constructions.

Triangulating these findings with the existing literature on the L1 acquisition of the structures at issue, we were able to conclude that the structures that child HSs found the most difficult were exactly the structures that usually emerge late (or very late) in monolingual language acquisition. This was also confirmed by a small scale study conducted on Portuguese monolingual children ranging in age between 12 and 13 years, based on the same cloze-test as the one administered to the bilinguals. Also for the monolinguals, relative pronouns, contracted prepositions and clitics were associated with the lowest accuracy scores.

Overall, these results suggest that the challenging structures for bilingual children represent a complex learning task also in monolingual language acquisition. In other words, child HSs acquire morphosyntactic structures through the same milestones as their monolingual peers, although they may lag behind in some linguistic domains that require more input to be successfully acquired. Notably, we also found that the accuracy in the use of these structures improved with age, highlighting a developmental trend among the bilinguals. In addition, these results do not sustain the assumption that CLI is the main factor contributing to developmental differences between heritage and monolingual children.

The main contribution of the present paper consists in showing that the difficulties exhibited by the bilingual children cannot be accounted for in terms of a single notion of complexity. Rather, different structures may be complex in acquisition for different reasons. In particular, we identified four main notions of complexity, as related to the different structures analysed in this contribution, i.e., derivational complexity, memory-based lexical forms, rules dependent on phonological or discourse contexts and multiple form-function mappings. In this sense, we moved away from the attempt to provide a single definition of complexity, but rather proposed a multifaceted view of this notion, which matches with extensive research on language acquisition.

References

Armon-Lotem, Sharon. 2005. The acquisition of subordination: From preconjunctionals to later use. In Dorit Ravid & Hava Bat-Zeev Shyldkrot (eds.), *Perspectives on language and language development*, 192–204. Dordrecht: Kluwer.

Armon-Lotem, Sharon, Ewa Haman, Kristine Jensen de López, Magdalena Smoczynska, Kazuko Yatsushiro, Marcin Szczerbinski, Angeliek van Hout, Ineta Dabašinskienė, Anna Gavarró, Erin Hobbs, Laura Kamandulytė-Merfeldienė, Napoleon Katsos, Sari Kunnari, Chrisa Nitsiou, Lone Sundahl Olsen, Xavier Parramon, Uli Sauerland, Reeli Torn-Leesik & Heather van der Lely. 2016. A large-scale cross-linguistic investigation of the acquisition of passive. *Language Acquisition* 23(1). 27–56. DOI: 10.1080/10489223.2015.1047095. (1080).

Batalha, Joana. 2018. *Relações entre o conhecimento explícito da língua e a competência de leitura*. Nova University of Lisbon. (Doctoral dissertation).

Bloom, Lois, Karin Lifter & Jeremie Hafitz. 1980. Semantics of verbs and the development of verb inflection in child language. *Language* 56. 386–412.

Bongartz, Christiane & Jacopo Torregrossa. 2020. The effects of balanced biliteracy on Greek-German bilingual children's secondary discourse ability. *International Journal of Bilingual Education and Bilingualism* 23. 948–963. DOI: 10.0050.2017.1355888. (1080).

Bowerman, Melissa. 1979. The acquisition of complex sentences. In P. Fletcher & M. Garman (eds.), *Language acquisition*, 285–306. Cambridge: Cambridge University Press.

Brito, Edvan P. 2018. The variation of obligatory preposition-article contractions in the interlanguage of adult learners of Portuguese. *Revista do GEL* 15. 241–256.

Catalão, M. Florentina Charneca. 2011. *Os padrões de uso dos pronomes pessoais àtonos em português europeu*. University of Évora. (MA thesis).

Chung, Eun Seong & Soojing Ahn. 2019. Examining cloze tests as a measure of linguistic complexity in L2 writing. *Language Research* 55. 627–649.

Clahsen, Harald, Fraibert Aveledo & Iggy Roca. 2002. The development of regular and irregular verb inflection in Spanish child language. *Journal of Child Language* 29(3). 591–622. DOI: 10.1017/S0305172.

Costa, João, Alexandra Fiéis & Maria Lobo. 2015. Input variability and late acquisition: Clitic misplacement in European Portuguese. *Lingua* 161. 10–26. DOI: 10.1016/j.lingua.2014.05.009.

Costa, Ana Luísa. 2006. *Complexidade estrutural de conectores concessivos*. Textos Selecionados do XXII Encontro da Associação Portuguesa de Linguística. Lisboa: APL.

Costa, João & Maria Lobo. 2007. Clitic omission, null objects or both in the acquisition of European Portuguese. In Sergio Baauw, Frank Drijkoningen & Manuela Pinto (eds.), *Romance Languages and Linguistic Theory 2005*, 59–72. Amsterdam: John Benjamins. DOI: 10.1075/cilt.291.06cos.

Costa, João & Maria Lobo. 2009. Clitic omission in the acquisition of European Portuguese: Data from comprehension. In Acrisio Pires & Jason Rothman (eds.), *Minimalist inquiries into child and adult language acquisition: Case studies across Portuguese*, 63–84. Berlin: Mouton de Gruyter. DOI: 10 . 1515 / 9783110215359.1.63.

Costa, João, Maria Lobo & Carolina Silva. 2009. Null objects and early pragmatics in the acquisition of European Portuguese. *Probus* 21(2). 143–162. DOI: 10.1515/prbs.2009.005.

Costa, João, Maria Lobo & Carolina Silva. 2012. Which category replaces an omitted clitic? The case of European Portuguese. In P. Larrañaga & P. Guijarro-Fuentes (eds.), *Pronouns and clitics in early language*, 105–129. Berlin: Mouton de Gruyter.

Costa, João, Maria Lobo & Carolina Silva. 2011. Subject–object asymmetries in the acquisition of Portuguese relative clauses: Adults vs. children. *Lingua* 121(6). 1083–1100. DOI: 10.1016/j.lingua.2011.02.001..

de Lourdes Gonçalves, Maria & Miriam Vinzentin. 2021. Portuguese in Switzerland: Investigating the representations of families and teachers. In A. Souza & S. Melo-Pfeifer (eds.), *Portuguese as a heritage language in Europe: A pluricentric perspective*, 119–150. Pontes Editores.

Diessel, Holger. 2004. *The acquisition of complex sentences*. Cambridge, UK: Cambridge University Press.

Dromi, Esther & Ruth A. Berman. 1986. Language-specific and language-general in developing syntax. *Journal of Child Language* 13. 371–381.

Estrela, Antónia. 2015. The acquisition of the passive structure in European Portuguese. *Journal of Child Language Acquisition and Development* 3(3). 153–168.

Flores, Cristina, Esther Rinke & Aldona Sopata. 2020. Acquiring the distribution of null and overt direct objects in European Portuguese. *Journal of Portuguese Linguistics* 19(5). 1–20. DOI: 10.5334/jpl.239.

Flores, Cristina, Ana Lúcia Santos, Alice Jesus & Rui Marques. 2017. Age and input effects in the acquisition of mood in heritage Portuguese. *Journal of Child Language* 44(4). 795–828. DOI: 10.1017/S030222.

Fox, Danny & Yosef Grodzinsky. 1998. Children's passive: A view from the by-phrase. *Linguistic Inquiry* 29(2). 311–332.

Hamann, Cornelia. 2006. Speculations about early syntax: The production of wh-questions by normally developing French children and French children with SLI. *Catalan Journal of Linguistics* 5. 143–189.

Jakubowicz, Celia. 2003. Computational complexity and the acquisition of functional categories by French-speaking children with SLI. *Linguistics* 41(2). 175–211. DOI: 10.1515/ling.2003.007.

Jakubowicz, Celia. 2011. Measuring derivational complexity: New evidence from typically developing and SLI learners of L1 French. *Lingua* 121(3). 339–351. DOI: 10.1016/j.lingua.2010.10.006.

Jakubowicz, Celia & Lea Nash. 2001. Functional categories and syntactic operations in (ab)normal language acquisition. *Brain and Language* 77. 321–339.

Jesus, Alice, Rui Marques & Ana Lúcia Santos. 2019. Semantic features in the acquisition of mood in European Portuguese. *Language Acquisition* 26(3). 302–338. DOI: 10.9223.2019.1570203. (1080).

Littlefield, Heather. 2009. *Evidence from English prepositions and their acquisition: A fine-grained approach to lexical and functional syntactic categories*. Riga, Latvia: VDM Verlag Dr. Müller.

Lobo, Maria & Carolina Silva. 2016. Ambiguidade pronominal em orações adverbiais do português europeu: Crianças vs. adultos. *Revista da Associação Portuguesa de Linguística* 2. 319–338.

Montrul, Silvina, Rebecca Foote & Silvia Perpiñán. 2008. Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning* 58. 503–553.

Montrul, Silvina & Tania Ionin. 2010. Transfer effects in the interpretation of definite articles by Spanish heritage speakers. *Bilingualism: Language and Cognition* 13(4). 449–473. DOI: 10.1017/S1366728910000040.

O'Grady, William, John Archibald, Mark Aronoff & Janie Rees-Miller. 2010. *Contemporary linguistics: An introduction*. 6th edn. Boston/New York: Bedford/St. Martin's.

Penner, Zvi. 1995. *Continuity constraints on paradigm formation in the acquisition of the complementizer system: A case study in Bernese Swiss German*. Unpublished manuscript. University of Bern.

Picoral, Adriana & Ana M. Carvalho. 2020. The acquisition of preposition + article contractions in L3 Portuguese among different L1-speaking learners: A variationist approach. *Languages* 5(4). DOI: 10.3390/languages5040045.

Polinsky, Maria. 2006. Incomplete acquisition: American Russian. *Journal of Slavic Linguistics* 14(2). 191–262.

Polinsky, Maria. 2008. Heritage language narratives. In Donna M. Brinton, Olga Kagan & Susan Bauckus (eds.), *Heritage language education: A new field emerging*, 149–164. London: Routledge.

Rauh, Gisa. 1993. On the grammar of lexical and non-lexical prepositions in English. In C. Zelinsky-Wibbelt (ed.), *The semantics of prepositions: From mental processing to natural language processing*, 99–150. Berlin: Mouton de Gruyter.

Rinke, Esther & Cristina Flores. 2014. Morphosyntactic knowledge of clitics by Portuguese heritage bilinguals. *Bilingualism: Language and Cognition* 17(4). 681–699. DOI: 10.1017/S136672891300076X.

Rinke, Esther & Cristina Flores. 2018. Another look at the interpretation of overt and null pronominal subjects in bilingual language acquisition: Heritage Portuguese in contact with German and Spanish. *Glossa* 3(1). 68. DOI: 10.5334/gjgl.535.

Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, MA: MIT Press.

Rizzi, Luigi. 2000. *Comparative syntax and language acquisition*. New York: Routledge.

Santos, Ana Lúcia. 2017. Alguns aspetos da aquisição de orações subordinadas completivas. In Maria João Freitas & Ana Lúcia Santos (eds.), *A aquisição de língua materna e não materna: Questões gerais e dados do português*, 249–273. Berlin: Language Science Press. DOI: 10.5281/zenodo.889436.

Santos, Ana Lúcia, Anabela Gonçalves & Nina Hyams. 2016. Aspects of the acquisition of object control and ECM-type verbs in European Portuguese. *Language Acquisition* 23(3). 199–233.

Santos, Ana Lúcia, Jason Rothman, Acrísio Pires & Inês Duarte. 2013. Early or late acquisition of inflected infinitives in European Portuguese? Evidence from spontaneous production data. In Misha Becker, John Grinstead & Jason Rothman (eds.), *Generative linguistics and acquisition: Studies in honor of Nina M. Hyams*, 65–88. Amsterdam: John Benjamins. DOI: 10.1075/lald.54.03san.

Santos, M. Fernanda Nascimento. 2002. *Os pronomes pessoais átonos no Português Europeu: Descrição de problemas que ocorrem no 3º ciclo e proposta de actividades didácticas*. University of Lisbon. (MA thesis).

Schneider, Phyllis, Rita Vis Dubé & Denyse Hayward. 2005. *The edmonton narrative norms instrument*. Retrieved from University of Alberta Faculty of Rehabilitation Medicine website. http://rehabresearch.ualberta.ca/enni/.

Schulz, Petra & Angela Grimm. 2019. The age factor revisited: Timing in acquisition interacts with age of onset in bilingual acquisition. *Frontiers in Psychology* 9. 10.3389/fpsyg.2018.02732, 2732.

Soares, Carla. 1998. *As categorias funcionais no processo de aquisição do Português europeu (estudo longitudinal da produção espontânea de uma criança de 1;2*. Universidade de Lisboa. (MA thesis).

Soares, Carla. 2003. The c-domain and the acquisition of European Portuguese: The case of wh-questions. *Probus* 15(1). 147–176. DOI: 10.1515/prbs.2003.002.

Sorace, Antonella. 2011. Pinning down the concept of "interface" in bilinguals. *Linguistic Approaches to Bilingualism* 1(1). 1–33. DOI: 10.1075/lab.1.1.01sor.

Sorace, Antonella, Ludovica Serratrice, Francesca Filiaci & Michela Baldo. 2009. Discourse conditions on subject pronoun realization: Testing the linguistic intuitions of older bilingual children. *Lingua* 119. 460–477.

Teodoro, Gabriel Malheiros. 2020. *A aquisição das preposições de, para e com por falantes do português do brasil.* Bachelor Dissertation, Universidade de Brasilia, Brazil.

Torregrossa, Jacopo, Maria Andreou, Christiane Bongartz & Ianthi M. Tsimpli. 2021. Bilingual acquisition of reference: The role of language experience, executive functions and cross-linguistic effects. *Bilingualism: Language and Cognition* 24(4). 694–706. DOI: 10.1017/S0000826.

Torregrossa, Jacopo, Christiane Bongartz & Ianthi M. Tsimpli. 2019. Bilingual reference production: A cognitive-computational account. *Linguistic Approaches to Bilingualism* 9(4-5). 569–599. DOI: 10.1075/lab.17026.tor.

Torregrossa, Jacopo, Cristina Flores & Esther Rinke. 2023. What modulates the acquisition of difficult structures in a heritage language? A study on Portuguese in contact with French, German and Italian. *Bilingualism: Language and Cognition* 26(1). 179–192.

Tsimpli, Ianthi. 2014. Early, late or very late? Timing acquisition and bilingualism. *Linguistic Approaches to Bilingualism* 4(3). 283–313.

Tuller, Laurice, Hélène Delage, Cécile Monjauze, Anne-Gaëlle Piller & Marie-Anne Barthez. 2011. Clitic pronoun production as a measure of atypical language development in French: a comparative study of SLI, mild-to-moderate deafness and benign epilepsy of childhood with centrotemporal spikes. *Lingua* 121(3). 423–441.

van Dijk, Chantal, Elise van Wonderen, Elly Koutamanis, Gerrit Jan Kootstra, Ton Dijkstra & Sharon Unsworth. 2022. Cross-linguistic influence in simultaneous and early sequential bilingual children: A meta-analysis. *Journal of Child Language* 49(5). 897–929. DOI: 10.1017/S0305000921000337.

van Riemsdijk, Henk. 1990. Functional preopositions. In Harm Pinkster & Inge Genee (eds.), *Unity in diversity: Papers presented to Simon C. Dik on his 50th birthday*, 229–242. Berlin: Mouton de Gruyter.

Varlokosta, Spyridoula, Adriana Belletti, João Costa, Naama Friedmann, Anna Gavarró, Kleanthes K. Grohmann, Maria Teresa Guasti, Laurice Tuller, Maria Lobo, Darinka Andjelkovic, Núria Argemí, Larisa Avram, Sanne Berends, Valentina Brunetto, Hélène Delage, Iris Ezeizabarrena Segurola María-José andFattal, Ewa Haman, Angeliek van Hout, Kristine M. Jensen de López, Napoleon Katsos, Lana Kologranic, Nadezda Krstić, Jelena Kuvac Kraljevic, Aneta Miękisz, Michaela Nerantzini, Clara Queraltó, Zeljana Radic, Sílvia Ruiz, Uli Sauerland, Anca Sevcenco, Magdalena Smoczyńska, Eleni Theodorou,

Alma van der Lely Heather andVeenstra, John Weston, Maya Yachini & Kazuko Yatsushiro. 2016. A cross-linguistic study of the acquisition of clitic and pronoun production. *Language Acquisition* 23(1). 1–26. DOI: 10.9223.2015.1028628.

Vasconcelos, Manuela. 1995. Relative clause acquisition and experimental research: A study with Portuguese children. In I.H. Faria & M.J. Freitas (eds.), *Studies on the acquisition of Portuguese*, 115–128. Lisboa: APL.Colibri.

Ximenes, Cristina. 2004. Preposition contraction in coordinated structures in Brazilian Portuguese. *MIT Working Papers in Linguistics* 47. 179–194.

Chapter 4

The complexity of word order change in a flexible system: On stability and variation in heritage Russian word order

Oksana Laleko
State University of New York at New Paltz

The story of heritage languages is often told as a story of grammatical simplification. A substantial body of work has pointed to numerous areas of reduced structural elaboration across heritage language systems, often manifested as a decrease in paradigmatic complexity through the elimination of sub-distinctions in various grammatical categories. Evidence for simplification in the domain of syntagmatic relations, such as in the encoding of information structure relations through word order, is somewhat weaker for heritage languages. This chapter examines the dynamics of heritage language word order change through the lens of Russian. I draw on data from a series of contextualized acceptability judgment tests with homeland Russian speakers and English-dominant heritage Russian speakers at two levels of proficiency, probing the distribution of canonical (SVO) and non-canonical word order structures, including subject-verb inversion (OVS) and object fronting (SOV, OSV). While underrating OVS orders, heritage speakers converge with the controls in their judgments of OSV and SOV sentences, with a trend toward overgeneralization of SOV in high-proficiency speakers. In terms of contextual licensing, focus-driven movement appears more stable than displacement related to givenness. These results offer two important implications. Broadly, they indicate that heritage speakers do not show an across-the-board preference toward SVO syntax, nor do they display a generalized difficulty with information structure marking, suggesting that heritage language word order change should not be viewed narrowly through the lens of simplification, linearization, and pragmatic unmarking. Second, they demonstrate that the facilitation of a non-canonical pattern in a heritage language word order system may occur independently of dominant language transfer effects, bringing into focus other driving forces of language change, including input frequency and universal constituent placement preferences rooted in basic cognitive and communicative principles.

Oksana Laleko. 2024. The complexity of word order change in a flexible system: On stability and variation in heritage Russian word order. In Maria Polinsky & Michael T. Putnam (eds.), *Formal approaches to complexity in heritage languages*, 73–99. Berlin: Language Science Press. DOI: ??

Oksana Laleko

1 Introduction

Within the burgeoning literature on heritage languages and their linguistic properties, two overarching and partially intertwined themes have recently come to dominate the scholarly landscape. On the one hand, many researchers have pondered whether or not the presently available empirical data on structural properties of heritage languages allow for broader generalizations bearing on the issue of heritage language typology. Is it in principle possible to identify features, processes, and patterns that would prove characteristic of heritage languages as a linguistic phenomenon, and if so, what is the nature of these properties and how would they vary along a continuum of heritage grammars, including its diachronic (i.e., acquisition and transmission) and synchronic (i.e., proficiency and individual variability) dimensions? While a number of researchers have suggested that such generalizations are conceivable – at the very least, for certain domains of heritage language design (Benmamoun et al. 2013, Lohndal et al. 2019, Polinsky & Scontras 2020), it is clear that much cross-linguistic work remains to be done on the empirical front in order to bring the existing proposals up to the highest levels of explanatory robustness through rigorous testing under typologically and sociolinguistically varied conditions.

The second theme, and one that the present paper takes as its primary impetus, concerns the long-standing conceptualization of heritage language change as a unidirectional process of a decrease in linguistic complexity. Building on traditions rooted in decades of work on language obsolescence and attrition studies (Dorian 1989, Sasse 2001), theoretical models of heritage language competence have drawn heavily on linguistic research that has attended most extensively to areas of reduced structural elaboration across heritage language systems. Setting aside for a moment issues related to heritage language processing, existing representational models have, accordingly, tended to construe heritage language grammars as smaller, more economically organized networks of structures with (i) an overall lower degree of paradigmatic complexity due to the elimination of sub-distinctions in grammatical categories (e.g., case, gender, aspect), and (ii) tighter and more locally defined syntagmatic relationships. As discussed recently in Laleko & Scontras (2021; see also other papers in the special issue of the *Heritage Language Journal* on heritage language complexity), few studies have systematically accounted for instances of complexity invariance, indicative of areas of stability in heritage language transmission, and even fewer have turned their attention to processes of complexification arising in heritage language systems. Yet, there is no a priori reason to expect complexity-preserving and even

complexity-increasing mechanisms of change to be inoperative in heritage language contexts. Since the blossoming of historical linguistic studies in the nineteenth century, the prevalence of these processes in language evolution has been well recognized (e.g., Lightfoot 1991), including their manifestations in situations of contact-induced change, of which heritage languages are a particular, and in fact very prominent, case.

Building on insights from typological and historical linguistic studies, one may turn attention to a number of language-internal and external factors to identify possible triggers of complexity-producing change in heritage languages. Elements of complexification could arise, for instance, as by-products of compensatory trade-offs across various linguistic sub-domains affected by selective loss of forms and subsequent redistribution of functions correlated with the retained forms. In the same vein, form-preserving and strengthening effects could stem from cross-linguistic influence on the heritage language of other languages spoken in the same linguistic niche, including most notably the societally dominant language and, a less-often explored possibility, other minority languages completing for the speakers' cognitive resources. Likewise, one could look for signs of complexification in the form of original innovations reflecting the unique socio-demographic and linguistic conditions in which heritage grammars are constructed, a line of inquiry that seems particularly promising for heritage languages developing within stable diasporic speech communities and transmitted intergenerationally. Whatever the extent and primary source(s) of complexification phenomena turn out to be in heritage languages, these general considerations warrant a more nuanced approach to modeling heritage language change, one in which the predicted structural outcomes would rest on a more explicit recognition of the multi-dimensional and multi-directional nature of language change as a linguistic process.

Taking as a point of departure the prevailing assumption that the story of heritage languages is largely a story of grammatical simplification, this chapter aims to expand the conception of heritage language change as a phenomenon of much broader scope. As one step towards this goal, I examine the dynamics of word order change in heritage Russian, drawing on data taken to represent two distinct stages of change from a cross-sectional standpoint, compared to the homeland baseline. As I detail further below, word order variation in Russian, along with other Slavic languages, provides a fitting opportunity for investigations of heritage language word order change through the lens of complexity change: As discourse-configurational systems with an underlying SVO typology, Slavic languages make heavy use of scrambling (Ross 1967) to encode information structure relations in discourse, yielding virtually every possible configuration

of clausal constituents possible and communicatively desirable under the right discourse-pragmatic conditions. Given that most theoretical studies converge on the assumption that non-canonical (i.e., scrambled) sentences are more complex in Russian than those following the canonical order (Sekerina 2003: 302), the study of variation in this domain of a heritage language system emerges as a *bona fide* case of complexity variation.

In order to position the above claim more directly within the existing models of linguistic complexity, we may wish to draw a further distinction between *absolute* and *relative* complexity, two conceptualizations of complexity that have been utilized, independently of each other or jointly, across a number of linguistic fields in an attempt to establish specific, practically useful metrics serving to demarcate what one would consider more or less complex in language structure and/or language use (see Laleko & Scontras 2021 for a recent discussion). Within this dichotomy, the former notion of absolute complexity has been operationalized most prominently in typological and diachronic studies focused on the structural characteristics of linguistic systems and subsystems, while the latter notion of relative complexity has gained much ground in psycholinguistic work and research on language acquisition, taking the standpoint of the language user as the key benchmark in assessing linguistic complexity.

Drawing on metrics advanced within each of these frameworks, and assuming SVO as the underlying basic order, one would only need to take a very broad glance at the existing theoretical analyses of scrambling in Slavic to observe that the derived status of discourse-dependent orders places them onto the higher end of the complexity scale relative to the canonical, base-generated order. In the absolute, system-centered sense, the increased complexity associated with discourse configurationality is suggested by a range of factors including, but not limited to: the added structural complexity of the projections hosting the scrambled components, the additional movement operations involved in their relocation, the overall greater number of rules that must be specified to license the occurrence of scrambled orders, the less transparent (i.e., less faithful) relationship between the underlying and surface representations, or, from a cross-linguistic perspective, their higher typological markedness.[1] From the relative, or user-based, perspective, adopted in many processing-based models of scrambling, non-canonical sentences come out as more complex as well: time and again, structures with displaced constituents have been shown to engage more computational and working memory resources than those following canonical orders (Gibson 1998, Just et al. 1996).

[1]The last observation does not hold for SOV orders (Greenberg 1963, Dryer 2013).

In light of these considerations, and under the conception of heritage language change as a shedding of complex structures, word order change in a flexible system should look like a process of word order levelling, manifested in the weakening and loss of non-canonical orders in favor of the default (here: least contextually-bound) pattern. In situations of a shared default between the contact languages, which is the case for the Russian-English dyad examined here, it is the predominant SVO order that emerges as a prime candidate for replacing the scrambled orders in the simplified system, predicting exponentially degraded ratings of non-SVO orders in both heritage speaker groups. Conversely, if heritage language word order change involves complexity-preserving mechanisms as part of its process, the affected grammars may retain or even make greater use of structures falling under umbrella of non-canonical orders, resulting in a distinct – but not necessarily universally simpler – system. To anticipate the presentation of the results that follows, it is the latter scenario that is borne out empirically.

The rest of the chapter is structured as follows. §2 brings into focus the relevant theoretical generalizations and empirical findings pertaining to word order variation in Russian. In §3, I offer a concise review of representative existing studies on word order change in heritage languages, highlighting findings that bear most closely on the phenomena addressed in this study. §4 presents the methodology and results of the study, while §5 discusses its findings and their implications.

2 Theoretical and experimental approaches to word order variation in Russian

At least since the writings of the Prague School linguists (Mathesius 1947, Firbas 1964), it has been recognized that the surface linearization of clausal constituents (and their parts) in Slavic languages, including Russian, is strongly regulated by the communicative principles related to the encoding of information structure. However, the exact nature of the encoded categories and the mechanism(s) of their impact on the linguistic structure remain subject to lively debates to this date. Broadly, analyses of word order variation in Russian may be conceived of as occupying a niche along the spectrum between two poles: functionalist approaches, striving to identify the elements of discourse and pragmatic structure correlated with word order variation, and syntactic approaches, focused on working out the linguistic principles involved in deriving the resulting structures. The

brief exposition of the literature in this section highlights the key insights gained from both perspectives.

Among the functionalist accounts, one of the central questions has been the problem of conceptualizing the relevant information-structural relations responsible for the flexibility of surface orders in Slavic. While converging on the observation that sentences in languages like Russian tend to be organized so as to allow for a predictable ordering between parts that are relatively less informative and parts that are relatively more informative, with the former preceding the latter, a number of dichotomies have been proposed to capture these relations. Pairs of terms such as *theme* and *rheme*, *topic* and *comment*, *givenness* and *newness*, *background* and *focus*, *presupposition* and *focus*, *topic* and *focus* have all been used for this purpose, sometimes interchangeably but often with important differences in theoretical assumptions (see Gundel & Fretheim 2004 for a terminological overview). In keeping with the general rule of linearly increasing informativeness as the primary guiding principle of sentence structuring in Russian (Gundel 1988), several scholars have further expanded the traditional binary partition into a tripartite template, in which the discourse-neutral information may be hosted medially between the two informationally-distinguished parts as illustrated in (1) below (Brun 2001, King 1995):

(1) (Topic) – (discourse-neutral information) – focus

In accordance with this structuring, constituents occurring at the right edge of the clause in Slavic are most straightforwardly interpreted as carrying new information, while constituents in earlier positions are more easily assigned a discourse-neural or given status. To facilitate such interpretation, constituents base-generated elsewhere in the clause – but associated with new information (i.e., presentational) *focus* – may surface at the right edge; conversely, constituents that are *given* (e.g., by virtue of having been mentioned in prior discourse or otherwise accessible) can shift away from their canonical right-edge positions, leaving other material in the focus domain. These scenarios are illustrated in the examples (2–4) below. In (2), the SVO sentence is presented in a broad-focus (i.e., "out of the blue") context, such that all of its constituents represent new information. In (3) and (4), presentational focus rests narrowly on the subject or the verb, respectively; all remaining constituents in the answer sentences are discourse-given by virtue of appearing in prior context. It should be noted that the movement involved in the non-canonical examples below is not of obligatory nature, making SVO orders possible across all contexts, with the focused constituent receiving prosodic stress *in situ* (see, e.g., Jasinskaja 2016).

(2) a. Kakie novosti?
what news-FEM.PL
'What's new?'

b. Papa prodal mašinu. SVO
dad-NOM.SG sell-PAST.SG car-ACC.SG
'Dad sold the car'

(3) a. Kto prodal mašinu?
who-NOM sell-PAST.SG car-ACC.SG
'Who sold the car?'

b. Mašinu prodal papa. OVS
car-ACC.SG sell-PAST.SG dad-NOM.SG
'Dad sold the car'

(4) a. Čto papa sdelal s mašinoj?
what dad-NOM.SG do-PAST.SG with car-ACC.SG
'What did dad do with the car?'

b. Papa mašinu prodal. SOV
dad-NOM.SG car.ACC.SG sell-PAST.SG
'Dad sold the car'

c. Mašinu papa prodal. OSV
car-ACC.SG dad-NOM.SG sell-PAST.SG
'Dad sold the car'

Within the syntactically-oriented strand of research on discourse-configurationality in Slavic, we once again find a spectrum of proposed models that vary in their assumptions and conclusions, ranging from those that bring the discourse requirements on scrambling directly into the syntax (King 1995) to those postulating a separate abstract level of representation, such as Functional Form (Bailyn 1995), Information-Structure (Junghanns & Zybatow 1997), or I-Structure (Kondrashova 1996), at which word order is linearized in accordance with the information-structural partition of the utterance. The nature and directionality of operations leading to such linearization are even less fully understood. Consensus has not been reached on such fundamental issues as whether the categories of topic and focus occupy dedicated syntactic projections (King 1995, Dyakonova 2009), are associated with features not linked to particular projections (Kondrashova 1996), or are split such that only topics, but not foci, project syntactically (Junghanns & Zybatow 1997); whether all movements are leftward

(Slioussar 2007) or are bi-directional (Sekerina 1997); and how the relevant operations are to be motivated (e.g., Titov 2020).

Despite these many points of disagreement, most existing analyses converge globally on the observation that while the surface ordering of constituents in Russian marks their information-structural roles, the underlying configuration of the Russian clause reflects the grammatical functions of the arguments, with SVO as the underlying, basic pattern (however, see King 1995 for arguments in favor of treating Russian as a VSO language in which subjects move out to serve as topics). This theoretical consensus is supported by experimental observations. Looking at the parsing strategies employed by Russian speakers in processing basic and scrambled constructions, Sekerina's (1997) eye-tracking study provided psycholinguistic evidence for SVO serving as the basic, discourse-neutral variant within the Russian word order system. Prior offline linguistic studies with monolingual Russian-speaking children and adults have reached the same conclusion. For example, in a repetition task reported in Bailyn (1995), Russian-speaking children between ages 3.8 and 5.5 demonstrated a much greater accuracy rates for SVO orders (83% correct) compared to non-SVO orders (40% correct). Studies with adults have similarly pointed to the predominance and functional versatility of the SVO pattern in Russian. Holden & Krupp (1987), for instance, conducted a written acceptability study in which participants were asked to rank the six possible orders appearing in neutral and discourse-specific contexts, assuming that all sentences are to be read with neutral intonation. Across all contexts, SVO emerged as the preferred pattern for Russian monolingual adults. In a more recent study with adult monolinguals, Kallestinova (2007) similarly found 98.9% of context-neutral transitive sentences to be produced in SVO form. Across various discourse contexts, the Russian speakers were overall found to favor SVO, OVS, and SOV orders in production, and while the remaining orders (VSO, VOS, and OSV) were not produced consistently, they were still recognized as acceptable on grammaticality judgment tasks (Kallestinova 2007).

Investigations of corpus data have generally been in line with the experimental findings discussed above; however, they have also unveiled substantial variation in the frequency of non-canonical orders in written and spoken Russian. According to a synopsis of prior corpus studies presented in Miller & Weinert (1998: 260), both modalities generally favor SVO as the predominant order, but with a quantitatively significant decrease in its occurrence in the spoken language: for (mostly) written Russian, SVO (79%), OVS (11%), OSV (4%), VOS (2%), SOV (1%), and VSO (1%); for spoken Russian, SVO (42%), SOV (34%), OSV (11%), and OVS (3%). A quick comparison of the distributional patterns for the remaining orders reveals a bias toward written registers for OVS structures and an association with

spoken language for object-fronted orders (SOV, OSV). In a more recent analysis of the Russian National Corpus, Billings (2015) reports the following word order frequencies based on a sample of 500 transitive sentences: SVO (89.6%), SOV (4.4%), OVS (2.4%), OSV (1.8%), VSO (1.6%), VOS (0.2%). While suggesting an overall more rigid sentence structure than one reported in prior studies, this pattern aligns well with the production results of Kallestinova's (2007) study, which identified the same three top orders for Russian. Since the majority of sentences in Billings's (2015) sample (74.4%) represent non-academic texts, the distribution of these orders by register, while suggestive of register-based variation, is presented in tentative terms. However, within the data for non-academic texts, comprising the majority of the sample, SVO, SOV, OVS, and OSV configurations emerge as the most productive orders (Billings 2015: 36). It is these four orders in Russian that are investigated in the present study.

3 Word order in heritage languages: A snapshot of current research

Whether or not heritage languages preserve the word order(s) generated in their corresponding baselines is a question that has received increasing attention from researchers in recent years, partly due to a spike in interest in the status of interface properties in bilinguals, triggered by several influential proposals, and partly as a result of a steadily improving access to heritage language corpora that has made such explorations possible (see Kisselev 2021). A substantial number of word order studies have turned the spotlight on syntactically-driven phenomena (e.g., Hopp & Putnam 2015, Kühl & Petersen 2018, Westergaard & Lohndal 2019 on V2, Anderssen & Westergaard 2020 on the interactions of pronominal argument placement with negation), finding these properties to be somewhat vulnerable but overall not drastically affected in the heritage varieties under investigation. In what follows, however, I will limit my literature survey to studies that have delved explicitly into the effects of information structure on word order variation in heritage language production and comprehension.

Within this cohort of studies, the encoding of the two key dimensions of information structure, focus and topicality (or givenness), have tended to be considered independently of each other, with the majority of published work focusing on the principles governing the encoding of the most informationally salient linguistic elements across heritage language systems. Such "focus on focus" bias has left significant gaps in our understanding of heritage language strategies related to the syntactic expression of givenness. Despite the asymmetry in the available

body of work, the overall results in both domains of inquiry have not reached a consistent level of uniformity across language dyads, proficiency levels, linguistic contexts examined, and experimental techniques employed for data collection, highlighting the need for a more synergistic evaluation of these factors in future studies.

With respect to the marking of focus, for example, even within a single narrowly defined domain, such as the distribution of SV and VS orders with intransitive verbs in relation to the informational prominence of the subject, one finds a considerable *concordia discors* across the scholarly terrain. Working with English-dominant heritage speakers, Zapata et al. (2005) and Laleko (2022) detected no consistent influence of focus, even in advanced speakers, on subject position with unergative predicates (which do not independently trigger VS) in heritage Spanish and Russian, respectively. However, Prada Pérez & Pascual y Cabo (2012) reported target-like contrasts in the same domain in English-dominant Spanish heritage speakers at three levels of proficiency, and van Osch & Sleeman (2018) documented an increase in the acceptability of the non-canonical VS orders in Dutch-dominant heritage Spanish speakers, an effect the authors attribute to dominant language influence. Expanding the scope of inquiry to transitive and di-transitive structures, studies by Hoot (2017) and Gómez Soler & Pascual y Cabo (2018) on Spanish in the U.S. reported strong effects of focus on subject and object placement in heritage speakers, a finding paralleled in Ionin et al. (2023) for U.S. Russian. At the same time, using processing measures, Sagarra et al. (2019) detected lower accuracy rates for OVS structures in a self-paced reading study of U.S. Spanish. Similarly, Hoot (2019) demonstrated target-like felicity judgments for constructions involving focus-movement in English-dominant heritage speakers of Hungarian, but the interpretations of these structures were found to differ from those in baseline speakers.

Little work has been done comparing the use of canonical and non-canonical orders across speakers of the same heritage language with different contact languages, but such studies can be pivotal in charting out the role of cross-linguistic influence on the development of the relevant constructions in the heritage language. Zuban et al. (2021) examined the production of SVO and OVS orders in heritage Russian in Germany and the U.S. and found a difference between German-dominant and English-dominant speakers, with a reduction in the use of OVS orders in the latter group. Earlier production studies of heritage Russian in the U.S. documented a similar quantitative decrease in the use of OVS and other non-canonical constructions in the oral narratives produced by heritage speakers (Isurin & Ivanova-Sullivan 2008, Ivanova-Sullivan 2014, Laleko & Dubinina 2018, Polinsky 2008) – a pattern also shown to hold in written essays produced

by heritage Russian learners (Kisselev 2019). Yet, in contrast to these college-age English-dominant heritage writers of Russian, their slightly younger German-dominant peers were found to maintain and even amplify the high flexibility of word order in Russian, with some evidence of a carry-over of the propensity for V-final orders to both main and subordinate clauses (Brehmer & Usanova 2015).

The effects of topicality or givenness on the placement of constituents have not been examined as extensively for heritage languages; still, a similarly non-uniform picture emerges from the examination of the presently available data, limited to date mostly to the study of topicalization and clitic dislocation phenomena in Spanish. Based on reduced occurrence and diverging judgements of topicalization and clitic left dislocation in heritage Spanish speakers, Zapata et al. (2005) argue for permeability of interface domains to cross-linguistic influence in bilinguals. Investigations of clitic left dislocations in Montrul (2010a,b) further point to significant proficiency effects in the acquisition of these structures in the heritage language. Conversely, looking at clitic right dislocation, a rare construction employed for topicalization, Leal et al. (2014) report no proficiency effects on the ratings of heritage Spanish speakers, who demonstrated target-like felicity contrasts in their judgments at intermediate and advanced levels of proficiency. More recently, Leal et al. (2015) and Sequeros-Valle et al. (2020) provide further experimental evidence of convergence between heritage and baseline Spanish speakers on topic constructions by examining clitic-doubled left dislocations, which bear an anaphoric relation to the discourse context. While overextending these structures to non-anaphoric contexts in acceptability judgments, heritage speakers are overall found to perform on par with baseline speakers in making context-appropriate distinctions in their use, with differences between groups attributed to task effects rather than to gaps in the underlying knowledge of the construction in heritage bilinguals.

In summary, despite growing attention of researchers to the issue of information structure marking through word order in heritage languages, the overall picture remains rather mixed, with much work remaining to be done to reach a more balanced assessment of the relevant phenomena across languages, for speakers at various proficiency levels, and via methodologically diverse techniques. As the brief synopsis above suggests, for Russian in particular, very little work has extended beyond oral or written production to examine constraints on the occurrence of basic and discourse-dependent orders in heritage speakers, and no studies have experimentally investigated parallels between the syntactic encoding of focus and givenness in speakers at distinct proficiency levels. This study seeks to accomplish this goal.

4 The study

The experimental data presented in this section come from a larger research project designed to investigate syntactic and prosodic realization of information structure in Russian as a heritage and second language. Here, I draw on three written contextualized acceptability judgment experiments to examine the effects of focus and givenness on constituent placement preferences in homeland and heritage language speakers. The role of focus is addressed in the first experiment, comparing ratings for SVO and OVS orders in two different information-structural configurations: under broad focus, as illustrated in example (2) above, and under narrow presentational focus on the subject, as shown in (3). The effects of givenness are the main focus of the second and third experiments, aimed to test the speakers' object placement preferences and targeting SOV and OSV constructions, respectively. Building on prior findings attesting to a correlation between VO/OV orders and object givenness in Russian (Sirotinina 1965, Sliousar 2007), ratings for sentences with post-verbal and preverbal accusative objects are obtained in two conditions: under broad focus, where the entire sentence, including the object constituent, carries new information, and under narrow focus on the verb, in which the object constituent represents given (i.e., old, known, shared, previously mentioned, accessible from prior discourse) information, as demonstrated in (4) above. Taking into account the frequently observed tendency for the pre-verbal Russian objects to be pronominalized (Miller & Weinert 1998: 260), the given-object condition in the second and third experiments is comprised of two sub-conditions, with nominal and pronominal objects analyzed separately from each other.

Considering the scarcity of controlled experimental data on word order preferences in heritage Russian, the main research questions of the study were formulated broadly: (i) overall, how do heritage speakers of Russian compare to homeland speakers in their syntactic encoding of focus and givenness; (ii) what effect does heritage language proficiency play in both domains? However, keeping in mind the issues addressed in this chapter, each of these broad questions has been operationalized into predictions that directly engage the problem of complexity in heritage language change. First, if word order restructuring proceeds in the direction of simplification, manifested as the amplification of the syntactic default shared by the bilinguals' word order grammars, we should observe a reduction in the acceptance of scrambled orders in the data from heritage speakers in comparison to the baseline. Second, taking as a starting point the traditional conception of heritage language proficiency as a linear continuum that displays greatest signs of restructuring at its lowest level and incrementally con-

verges with the baseline at its highest level (cf. the creole continuum model in Bickerton 1975 and its reconceptualization for heritage languages in Polinsky & Kagan 2007), the word order system instantiated in the grammars of the higher-proficiency speakers should be closer to the baseline system than one constructed by speakers at a lower point on the continuum.

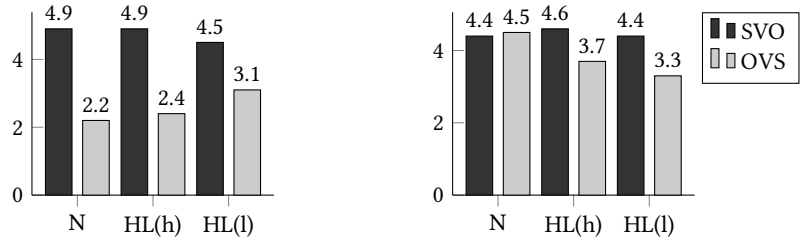
To test these predictions, acceptability ratings were gathered from forty-two adult speakers of Russian, including twenty-seven English-dominant heritage speakers (mean age = 19.4) and fifteen monolingually raised Russian-speaking controls (mean age = 24). All heritage speakers identified English as their primary and most frequently used language of daily communication and designated Russian as their non-dominant and less frequently used language (mean of average daily use = 23.7%). Consistent with the canonical profile of early imbalanced naturalistic bilinguals (Aalberse et al. 2019, Montrul 2016, Polinsky 2018), all heritage speakers in the study reported hearing and speaking Russian in early childhood, followed by an increase in the use of English coinciding, on average, with the age of entering the public school system (mean age of switch to English = 4.5). In contrast, all speakers in the control group reported Russian as the only language used in all daily communication and provided at-ceiling proficiency self-ratings across the four key areas of linguistic competence (understanding, speaking, reading, and writing Russian). The heritage speakers' self-assessments of the same four areas revealed the expected imbalance between skills acquired in a naturalistic environment (mean rating = 8.2 on a 10-point scale for understanding, mean rating = 7.1 for speaking) and those associated with explicit, formalized learning (mean rating = 6.5 for reading, mean rating = 6 for writing), suggesting a direct correlation between the speakers' degree of confidence in their linguistic abilities and the extent of their exposure to contexts in which these abilities are typically acquired.

In addition to the self-ratings of the Russian language proficiency obtained as part of the socio-demographic questionnaire, all participants completed an independent proficiency test designed to assess the speakers' morpho-syntactic knowledge across several grammatical domains previously shown to undergo restructuring and correlate with proficiency in heritage language speakers (Polinsky 2018). The test consisted of ten sentences containing violations in the areas of grammatical gender, subject-verb agreement, and pro-drop. On the basis of the obtained proficiency scores (for monolinguals, the mean score was predictably low at 1.6 on a 1–5 Likert scale), the heritage speakers were divided into two groups, yielding a higher-proficiency group of 12 speakers (mean score of 1.8) and a lower-proficiency group of 15 speakers (mean score of 2.8), with a statistically significant difference obtained between the two heritage speaker groups

based on proficiency ($t(22) = -7.2$, $p < 0.01$). Participants with a mean proficiency score of 4.0 or above were not included in the study.

The speakers rated 48 target sentences, intermixed with 60 fillers and presented in a randomized manner in Experigen (Becker & Levine 2013). The stimuli were presented to the participants as question-answer sequences, with one question followed by two answer choices with a 5-point Likert scale appearing next to each of the two answers. This design was deemed optimal to ensure that any differences in the participants' ratings of the relevant word order configurations reflected their judgments of word order proper, keeping the lexical and morphological information identical between the canonical and non-canonical sentences being evaluated.

The means for the relevant conditions of the study, presented in Figures 1–3 below, were analyzed statistically using Welch's unequal variances *t*-tests. First, I present the results on the syntactic encoding of focus in the three varieties[2] of Russian under investigation, comparing SVO and OVS sentences under broad focus and under narrow focus on the subject constituent (Figure 1).



(a) Broad (b) Narrow

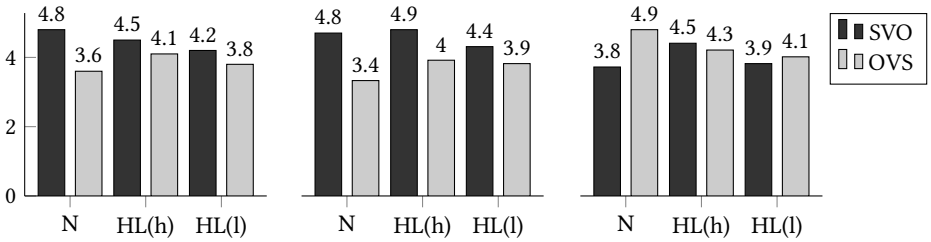
Figure 1: Ratings for SVO and OVS sentences in all-new-information and subject-focus contexts

In line with the theoretical generalizations about the unmarked, contextually unrestricted status of SVO in Russian, all participants demonstrated very high acceptance rates for SVO orders across all experimental conditions. Furthermore, both the homeland speakers and the higher-proficiency heritage speakers displayed nuanced judgments of OVS orders based on information structure, rating these structures significantly higher under narrow focus than under broad focus: controls ($t(82.2) = -10.3$, $p < 0.01$); HL(h) ($t(51.8) = -3.7$, $p < 0.01$). These results indicate that the interaction of OVS orders and subject focus, frequently

[2]In all figures, the data are presented in the following order: N = homeland speakers, HL(h) = higher-proficiency heritage speakers, HL(l) = lower-proficiency heritage speakers.

discussed in the theoretical literature on Slavic and confirmed experimentally in this study, is actively operative at high levels of heritage language proficiency. In contrast, the lower-proficiency heritage speakers rated OVS orders similarly in broad-focus and narrow-focus conditions: $t(61.2) = -0.4$, $p > 0.05$), suggesting no sensitivity to this information-structural constraint in these speakers' grammars. While attesting to baseline-like principles on the occurrence of OVS orders in high-proficiency speakers, the results nevertheless reveal a quantitative reduction in the acceptability of these orders by heritage speakers at both levels of proficiency, compared to the homeland speakers. Thus, in the narrow focus condition: controls vs. HL(h) ($t(63.08) = 3.01$, $p < 0.01$); controls vs. HL(l) ($t(76.8) = 4.46$, $p < 0.01$).

The remaining data bear on the encoding of givenness in the varieties of Russian examined. First, I outline the results obtained for SVO and SOV sentences presented in all-new-information contexts and in contexts that identify the object constituent as discourse-given. In the latter condition, data for sentences with nominal and pronominal objects are presented separately (Figure 2).



(a) All-new (b) Object-given (c) Object-given (pron.)

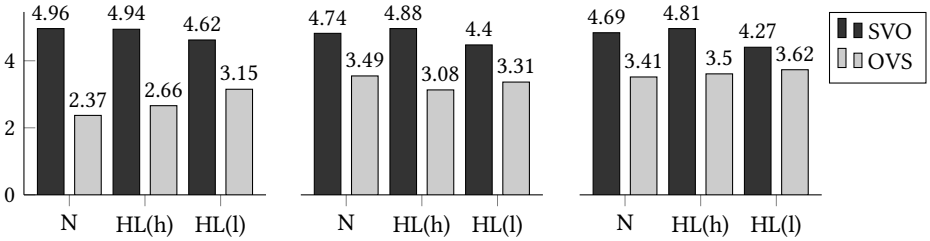
Figure 2: Ratings for SVO and SOV sentences in all-new-information and object-given contexts

In all-new contexts, Russian speakers in the control group displayed a significant preference towards SVO over SOV orders: $t(52.9) = 5.04$, $p < 0.01$. However, heritage speakers in both groups showed no such preference, accepting both SVO and SOV equally under broad focus (HL(h): $t(39.2) = 1.32$, $p > 0.05$; HL(l): $t(58) = 1.27$, $p > 0.05$). In object-given contexts with non-pronominal objects, the higher-proficiency heritage speakers and baseline controls displayed a preference for SVO over SOV orders: controls ($t(68.3) = 7.67$, $p < 0.01$), HL(h) ($t(40.6) = 4.16$, $p < 0.01$), while the lower-proficiency heritage speakers continued to treat SVO and SOV structures as interchangeable: $t(86.7) = 1.79$, $p > 0.05$. In pronominal constructions, however, the homeland speakers strongly preferred SOV orders to SVO orders ($t(62.9) = -5.86$, $p < 0.01$), while the heritage speakers found both

options as equally acceptable, regardless of proficiency (HL(h): $t(59.06) = 0.72$, $p > 0.05$; HL(l): $t(86.97) = -0.75$, $p > 0.05$).

On across-group comparisons, SOV orders received significantly higher ratings in the higher-proficiency heritage language group compared to the baseline group both in all-new contexts ($t(48.1) = -1.6$, $p < 0.05$) and in object-given contexts ($t(73.8) = -2.04$, $p < 0.05$). Speakers in the lower-proficiency heritage speaker group converged with the controls in both conditions: all-new ($t(61.2) = -0.69$, $0 > 0.05$), object-given ($t(87.8) = -1.73$, $p > 0.05$). Only in the pronominal condition, associated with an SOV preference over SVO in homeland speakers, did their ratings of SOV exceed those of the bilingual speakers: controls vs. HL(h) ($t(41.8) = 2.83$, $p < 0.01$), controls vs. HL(l) ($t(52.1) = 3.83$, $p < 0.01$).

With respect to factors conditioning the SVO/SOV alternation in Russian, object givenness did not prove to be a crucial determinant in either homeland or heritage speakers at either level of proficiency: in all three groups, SOV orders were ranked uniformly in statistical terms regardless of their occurrence in all-new or object-given sentences: baseline ($t(68.9) = 0.58$, $p > 0.05$), HL(h) ($t(47.6) = 0.45$, $p > 0.05$), HL(l) ($t(67.2) = -0.3$, $p > 0.05$). However, ratings for SOV constructions were strongly affected by object type in homeland speakers, who significantly favored SOV orders with pronominal objects over non-pronominal objects: $t(62.9) = -8.2$, $p < 0.01$. Neither group of heritage speakers demonstrated significant effects with respect to object type (HL(h): $t(69.8) = -1.18$, $p > 0.05$; HL(l): $t(87.9) = -0.76$, $p > 0.05$).



(a) All-new (b) Object-given (c) Object-given (pron.)

Figure 3: Ratings for SVO and OSV sentences in all-new-information and object-given contexts

Finally, I turn to the analysis of OSV constructions. All groups showed a robust global preference towards SVO over OSV orders, both in all-new contexts (controls: $t(52.33) = 14.1$, $p < 0.01$; HL(h): $t(38.48) = 8.93$, $p < 0.01$; HL(l): $t(75.4) = 5.71$, $p < 0.01$) and in object-given nominal contexts (controls: $t(70.20) = 5.68$, $p < 0.01$); HL(h): $t(40.17) = 7.15$, $p < 0.01$; HL(l): $t(74.62) = 3.97$, $p < 0.01$). With

pronominal objects, this tendency was observed in baseline and high-proficiency heritage speakers: controls ($t(84.02) = 5.33$, $p < 0.01$), HL(h) ($t(42.24) = 5.16$, $p < 0.01$). In lower-proficiency speakers, the difference between SVO and OSV orders with given pronominal objects did not reach significance at the same threshold, but was still significant at a weaker level: $t(83.14) = -1.51$, $p < 0.05$.

Despite exhibiting a preference for the canonical SVO pattern over OSV in all conditions, the homeland speakers were highly sensitive to object givenness in their ratings of OSV orders and gave significantly higher ratings to OSV in object-given contexts compared to all-new contexts: $t(98.77) = -4.07$, $p < 0.01$. In contrast, the heritage speakers were not sensitive to object givenness in their ratings of OSV orders at either proficiency level: HL(h) ($t(69.96) = -1.19$, $p > 0.05$, HL(l) ($t(87.57) = -0.49$, $p > 0.05$). Furthermore, object type (i.e., pronominal or not) did not matter for any of the three groups: homeland ($t(99.99) = 0.27$, $p > 0.05$), HL(h) ($t(69.99) = -1.21$, $p > 0.05$, HL(l) ($t(87.14) = -0.99$, $p > 0.05$). In what follows, I discuss these results and their implications.

5 Discussion and conclusions

The study examined the dynamics of word order change in heritage Russian, with the larger aim of assessing to what extent the trajectory of change in this domain conforms to the broad conception of heritage language change as successive simplification of structures and structure-building processes along the proficiency scale, emerging from a rich body of previous work on morphosyntactic restructuring in heritage languages. Experimental data from heritage Russian speakers at two levels of proficiency and a group of homeland speakers were analyzed to identify principles utilized in the respective linguistic varieties for the syntactic encoding of information-structural categories: focus, involved in the alternation between SVO and OVS orders, and givenness, implicated in the variation among SVO, SOV, and OSV orders. Starting with the theoretically established premise that Russian builds on an SVO tree to derive OVS, SOV, and OSV structures, and that these structures occur in particular, predetermined discourse contexts, the main goal of the study was to gauge the stability of the derived orders across two levels of heritage language proficiency and to trace changes in principles guiding their occurrence.

Broadly, the results do not support the conceptualization of heritage language change as a process of inexorable simplification. Instead, the findings point to a more nuanced picture of word order change, whereby patterns of simplification in some domains of the word order system intersect with trends towards preservation and even increase of complexity in its other areas. In the discussion of the

results that follows, I will first outline and comment on the manifestations of reductive change, and then take stock of areas of stability and amplified variability in the heritage language word order system.

With respect to the OVS constructions, employed for the encoding of subject focus in Russian, the trajectory of change emerging from the ratings of homeland speakers and heritage speakers at two discrete points on the proficiency spectrum renders support to a model of heritage language change as a linearly progressing process of reduction in syntactic variability. While heritage speakers at the higher end of the proficiency spectrum display the same grammaticality effects as homeland speakers, these effects dissipate with decreasing proficiency. Furthermore, heritage speakers in both groups display a diminished acceptance of OVS, preferring instead the canonical SVO structures in the same contexts. These patterns are indicative of a weakening of OVS in the heritage language and its replacement with the canonical SVO pattern in contexts in which both structures co-exist in the baseline system. In effect, this change may be viewed as a *narrowing* of a set of linguistic options available for the marking of a particular distinction in favor of the least restrictive, minimally specified variant that constitutes a shared default between the bilinguals' two word order systems.

Data on the distribution of OSV orders in Russian present a less clear case of simplification-driven change in the heritage language word order system. Since no quantitative reduction in the acceptability of these orders was observed – either in the group of heritage speakers compared to homeland speakers, or as a function of heritage language proficiency – it would be premature to conclude that these non-canonical structures are on a path to elimination from the heritage Russian word order system. While showing no immediate signs of quantitative decrease, the results are nevertheless indicative of a change affecting the contextual principles on the occurrence of OSV orders in the heritage language: the robust effects of object givenness attested in the homeland variety are absent in the heritage speakers' ratings regardless of proficiency. In this sense, the findings partially mirror the pattern of change observed for OVS structures, with an important difference concerning the proficiency level at which the loss of the relevant information-structural effect is observed. Recall that the deactivation of focus as a conditioning factor for OVS structures was attested only in the lower-proficiency heritage group, with advanced speakers performing on par with homeland speakers on this dimension. In contrast, the unlinking of OSV from object givenness is discernible at both proficiency levels. This contrast highlights a previously observed asymmetry between the effects of two distinct facets of information structure on word order, whereby the more informationally and prosodically salient categories like focus and contrast seem more re-

silient to change than categories associated with topicality and anaphoricity (see Laleko 2021: 719–721 for discussion). If on the right track, these results advocate for a more fine-grained approach to the study of the syntax-discourse interface phenomena than one commonly used in empirical research, where "information structure" often stands as a blanket term rather than a composite, multi-layered category comprised of an array of phenomena with distinct effects on linguistic structure and interpretation.

Finally, the patterns of homeland and heritage speakers' ratings obtained for SOV constructions provide the strongest evidence against an overly broad conceptualization of heritage language word order change as a unidirectional process of convergence towards the SVO template. Even at the lower proficiency level, heritage speakers showed no decrease in acceptability for SOV orders with nominal objects, fully converging with homeland speakers on SOV constructions in both all-new and object-given contexts – a pattern indicative of their high stability in bilingual grammars. Moreover, the fact that the higher-proficiency heritage speakers displayed more favorable judgments of SOV orders than homeland speakers in the same contexts speaks not only to stability, but also to an enhancement and strengthening of this pattern in these more established heritage grammars. The fact that the higher-proficiency speakers in this study rated SOV structures on par with SVO structures in all-new contexts is particularly notable in this regard: in effect, this development constitutes an expansion in the set of defaults in the heritage language word order system, compared to the baseline system. This pattern of change stands in sharp contrast with the view of heritage language change as a process of contraction to a single default (e.g., of the type observed in this study for SVO/OVS alternation), suggesting that multiple defaults are, in principle, permitted in restructured grammars and prompting further research into factors responsible for their emergence. In the remainder of this concluding chapter, I frame the discussion around two particular implications of these results, concentrating primarily on avenues in which they may inform future research on heritage language change.

Viewed in the context of the bulk of existing studies, the observed "elevation" of the SOV pattern to the status of unmarked, discourse-neutral order in the grammars of English-dominant Russian speakers constitutes a rarely documented case of non-transfer-induced complexity-preserving change. While several prior heritage language studies have reported evidence of complexification in various domains of heritage language structure, manifested as a strengthening of a particular non-canonical order or construction in a heritage language variety otherwise undergoing reductive change (e.g., Aalberse & Moro 2014, Aalberse & Andringa 2017, Brehmer & Usanova 2015, van Osch & Sleeman 2018), virtually all

of such hitherto reported developments have been attributed to cross-linguistic influence from the societally dominant language. In this light, the observed retention and strengthening of SOV, a structure that is impossible in English, in heritage Russian provides a strong impetus for continued research on sources of complexity-preserving change operating in contact situations but falling outside of clear transfer effects. The present results bring two such factors into the spotlight – one having to do with the internal composition of the Russian word order system, and the other stemming more globally from the cognitive and communicative principles that govern constituent placement in natural languages. Focusing narrowly on the internal dynamic of the Russian language, the marked contrast between the frequencies of VO and OV orders observed for its written/formal and spoken/colloquial registers, with preverbal objects ranging between 7–9% in scientific speech and up to 60% in colloquial speech, have prompted some scholars to characterize spoken Russian as undergoing a typological shift from SVO to SOV (Slioussar 2007). Under this view, the expansion of the SOV pattern in heritage Russian may be seen as an instance of input-driven change that builds on and amplifies incipient changes already happening in the baseline (Polinsky 2018); the fact that this development in the heritage language occurs despite the potentially constraining effects of ambient language transfer and seems resistant to them attests to its potency as a mechanism of change.[3] From a more general typological perspective, the pattern of SOV strengthening in spoken Russian, including Russian as a heritage language, may be a reflection of the basic evolutionary and cognitive principles that account for the overall preference for SOV ordering in human language (Dryer 2013, Newmayer 2000) and have been claimed to be engaged productively in the formation of emerging linguistic systems and in spontaneous communication, including under conditions of limited input (Goldin-Meadow 2003, Napoli & Sutton-Spence 2014). If on the right track, these proposals warrant a closer look at the fate of SOV structures in heritage languages – systems modeled on reduced and variably accessible input. Such investigations would be particularly welcome for languages with an otherwise restricted occurrence of SOV, e.g. heritage English, as a way of teasing apart the issue of universal constituent placement principles from language-internal, diachronically changing pressures – a task that cannot be accomplished in this study.

[3]It should be noted that input-related properties may also be implicated in the observed quantitative reduction of OVS structures to mark subject focus in heritage Russian. It has been observed that colloquial Russian differs from standard Russian in the syntactic encoding of new information focus: while usually expressed clause-finally in written language, focus is often preposed in spoken registers (Krylova & Khavronina 1986, Yokoyama 1986).

A second, related implication emerging from the patterns of results obtained for SOV structures concerns the construct of heritage language proficiency and its conceptualization and operationalization in heritage language research. Since the inception of the field, the study of heritage language properties has been rather heavily toward morphosyntax, and models of language proficiency proposed to account for the systematic variability in the linguistic abilities in heritage speakers, most typically scaled from low to high proficiency based on a degree of distance from the baseline system, have been calibrated first and foremost to aspects of grammatical variation, with a tacit assumption that other domains of language would follow suit. The results presented here do not support this assumption, suggesting instead that the extent of convergence between the heritage and homeland varieties may vary, and sometimes principally, for individual linguistic domains. In this study, which relied on morphosyntactic variables for proficiency assessment, heritage systems with deeper grammatical restructuring were more baseline-like in their word order properties than those with more stable grammars. While counter-intuitive under traditional conceptions of heritage language proficiency, such diverging trajectories of change across different linguistic sub-domains – morphosyntax and word order in particular – have been documented in other studies of language shift. For instance, a similar pattern of proficiency split was attested in shifting speakers of Even, a fixed head-final SOV language spoken in northeastern Russia. The less proficient speakers resisted word order change while omitting inflectional morphology; in contrast, the higher-proficiency speakers exhibited word order changes while maintaining the grammar of case (Kantarovich et al. 2021). Taken together with other studies that document innovations in stable and successfully maintained heritage languages (Aalberse & Moro 2014, Aalberse & Andringa 2017), the results obtained here raise the question of whether morphosyntactic stability may serve as a precursor of change in language domains that are especially responsive to communicative pressures. One prediction of this hypothesis would be that the more established heritage grammars, as measured by grammatical accuracy, may be associated with greater heritage language use and exposure, leading in turn to a higher degree of innovations arising as a result of the language developing in a new environment and under different communicative conditions.

If on the right track, this observation encourages a closer look at proficiency-based variation in heritage languages as an opportunity to expand our conceptualization of heritage language change in order to arrive at a more fine-grained differentiation among distinct mechanisms contributing to its genesis, teasing apart elements of change arising due to quantitative reduction in the linguistic input from those stemming from propagation of features and norms entrenched

in qualitatively different, yet abundantly present, input. With respect to the data discussed here, it appears that while elements of simplification are certainly a part of the process of change, they are not, at this in this case, the only relevant dynamic of change in the heritage language varieties under examination.

References

Aalberse, Suzanne & Francesca Moro. 2014. Stabilility in Chinese and Malay heritage languages as a source of divergence. In Kurt Braunmüller, Stefan Höder & Karoline Kühl (eds.), *Stability and divergence in language contact: Factors and mechanisms* (Studies in Language Variation 16), 141–162. Amsterdam: John Benjamins. DOI: 10.1075/silv.16.07aal.

Aalberse, Szanne, Ad Backus & Pieter Muysken. 2019. *Heritage languages: A language contact approach.* Amsterdam/Philadelphia, PA: John Benjamins. DOI: 10.1075/sibil.58.

Aalberse, Yiwen, Suzanne andd Zhou & Sible Andringa. 2017. Extended use of demonstrative pronouns in two generations of Mandarin Chinese speakers in the Netherlands: Evidence of convergence? In Elma Blom, Lleonie Cornips & Jeannette Schaeffer (eds.), *Cross-linguistic influence of bilingualism*, 25–48. Amsterdam: John Benjamins.

Anderssen, Merete & Marit Westergaard. 2020. Word order variation in heritage language: Subject shift and object shift in Norwegian. In Bernhard Brehmer & Jeanine Treffers-Daller (eds.), *Lost in transmission: The role of attrition and input in heritage language development* (Studies in Bilingualism 59), 99–124. Amsterdam: John Benjamins.

Bailyn, Jonathan. 1995. *A configurational approach to Russian "free" word order.* Ithaca, NY: Cornell University. (Doctoral dissertation).

Becker, Michael & Jonathan Levine. 2013. *Experigen: An online experiment platform.* Available at http://becker.phonologist.org/experigen.

Benmamoun, Elabbas, Silvina Montrul & Maria Polinsky. 2013. Heritage languages and their speakers: Opportunities and challenges for linguistics. *Theoretical Linguistics* 39(3-4). 129–181.

Bickerton, Derek. 1975. *Dynamics of a creole continuum.* Cambridge, UK: Cambridge University Press.

Billings, Stephanie K. 2015. *A corpus-based analysis of Russian word order patterns.* Brigham Young University. (MA thesis).

Brehmer, Bernhard & Irina Usanova. 2015. Let's fix it? Cross-linguistic influence in word order patterns of Russian heritage speakers in Germany. In Hagen Peukert (ed.), *Transfer effects in multilingual language development* (Hamburg Studies on Linguistic Diversity 4), 161–188. Amsterdam: John Benjamins.

Brun, Dina. 2001. Information structure and the status of NP in Russian. *Theoretical Linguistics* 27. 109–135.

Dorian, Nancy. 1989. *Investigating obsolescence: Studies in language contraction and death*. Cambridge: Cambridge University Press.

Dryer, Matthew S. 2013. Order of subject, object and verb. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/81 (7 March, 2022).

Dyakonova, Marina. 2009. *A phase-based approach to Russian free word order*. Amsterdam, Netherlands: University of Amsterdam. (Doctoral dissertation).

Firbas, Jan. 1964. On defining the theme in functional sentence analysis. *Travaux Linguistique de Prague* 1. 267–280.

Gibson, Edward. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68. 1–76.

Goldin-Meadow, Susan. 2003. *The resilience of language: What gesture creation in deaf children can tell us about how all children learn language*. New York, NY: Psychology Press.

Gómez Soler, Inmaculada & Diego Pascual y Cabo. 2018. On focus and weight in Spanish as a heritage language. *Spanish Journal of Applied Linguistics* 31(2). 437–466.

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of Language*, 73–113. Cambridge, MA: The MIT Press.

Gundel, Jeanette K. 1988. *The role of topic and comment in linguistic theory*. New York: Garland.

Gundel, Jeanette K. & Thorstein Fretheim. 2004. Topic and focus. In Laurence R. Horn & Gregory Ward (eds.), *Handbook of pragmatics*, 175–196. Oxford: Blackwell Publishing.

Holden, Kyril T. & Maria Krupp. 1987. Word order in Russian transitive sentences. *Folia Slavika* 8. 254–271.

Hoot, Bradley. 2017. Narrow presentational focus in heritage Spanish and the syntax-discourse interface. *Linguistic Approaches to Bilingualism* 7(1). 63–95.

Hoot, Bradley. 2019. Focus in heritage Hungarian. *Language Acquisition* 26(1). 46–72.

Hopp, Holger & Michael T. Putnam. 2015. Syntactic restructuring in heritage grammars: Word order variation in Moundridge Schweitzer German. *Linguistic Approaches to Bilingualism* 5(2). 180–213. DOI: 10.1075/lab.5.2.02hop.

Ionin, Tania, Maria Goldshtein, Tatiana Luchkina & Sofya Styrina. 2023. Who did what to whom, and what did we already know? *Linguistic Approaches to Bilingualism* 13(3). 343–371.

Isurin, Ludmila & Tanya Ivanova-Sullivan. 2008. Lost in between: The case of Russian heritage speakers. *Heritage Language Journal* 6(1). 72–104.

Ivanova-Sullivan, Tanya. 2014. *Theoretical and experimental aspects of syntax-discourse interface in heritage grammars*. Leiden: Brill.

Jasinskaja, Katja. 2016. Information structure in Slavic. In Caroline Féry & Shinichiro Ishihara (eds.), *The Oxford handbook of information structure*, 709–732. Oxford, UK: Oxford University Press.

Junghanns, Uwe & Gerhild Zybatow. 1997. Syntax and information structure of russian clauses. In *Formal approaches to slavic linguistics 4: the cornell meeting*. 289–319. Ann Arbor, MI: Michigan Slavic Publications.

Just, Marcel, Patricia A. Carpenter & Darold D. Hemphill. 1996. Constraints on processing capacity: Architectural or implementational? In *Mind matters: A tribute to Allen Newell*, 141–178. Mahwah, NJ: Lawrence Erlbaum Associates.

Kallestinova, Elena D. 2007. *Aspects of word order in Russian*. University of Iowa. (Doctoral dissertation).

Kantarovich, Jessica, Lenore A. Grenoble, Antonia Vinokurova & Elena Nesterova. 2021. Complexity and simplification in language shift. *Frontiers in Communication* 6. 1–17. DOI: 10.3389/fcomm.2021.638118.

King, Tracy H. 1995. *Configuring topic and focus in Russian*. CSLI Publications.

Kisselev, Olesya. 2019. Word order patterns in the writing of heritage and second language learners of Russian. *Russian Language Journal* 69. 149–174.

Kisselev, Olesya. 2021. Corpus-based methodologies in the study of heritage languages. In Silvina Montrul & Maria Polinsky (eds.), *The Cambridge handbook of heritage languages and linguistics*, 520–544. Cambridge, UK: Cambridge University Press.

Kondrashova, Natalia Y. 1996. *The syntax of existential quantification*. University of Wisconsin-Madison. (Doctoral dissertation).

Krylova, O. A. & S. A. Khavronina. 1986. Moscow: Russkij Jazik.

Kühl, Karoline & Jan H. Petersen. 2018. Word order in American Danish declaratives with a non-subject initial constituent. *Journal of Language Contact* 11. 413–440.

Laleko, Oksana. 2021. Discourse and information structure in heritage languages. In Silvina Montrul & Maria Polinsky (eds.), *The Cambridge handbook of heritage languages and linguistics*, 691–727. Cambridge, UK: Cambridge University Press.

Laleko, Oksana. 2022. Word order and information structure in heritage and L2 Russian: Focus and unaccusativity effects in subject inversion. *International Journal of Bilingualism* 26(6). 749–766.

Laleko, Oksana & Irina Dubinina. 2018. Word order production in heritage Russian: Perspectives from linguistics and pedagogy. In Susan Bauckus & Susan Kresin (eds.), *Connecting across languages and cultures: A heritage language festschrift in honor of olga kagan*, 191–215. Bloomington, IN: Slavica.

Laleko, Oksana & Gregory Scontras. 2021. On the many dimensions of complexity in hertiage languages. *Heritage Language Journal* 18(2). 1–37. DOI: 10.1163/15507076-12340009.

Leal, Tanja, Jason Rothman & Roumyana Slabakova. 2014. A rare structure at the syntax-discourse interface: Heritage and Spanish-dominant native speakers weigh in. *Language Acquisition* 21(4). 411–429.

Leal, Tanja, Jason Rothman & Roumyana Slabakova. 2015. Discourse-sensitive clitic-doubled dislocations in heritage Spanish. *Lingua* 155. 85–97.

Lightfoot, David. 1991. *How new languages emerge*. Cambridge, UK: Cambridge University Press.

Lohndal, Terje, Jason Rothman, Tanja Kupisch & Marit Westergaard. 2019. Heritage language acquisition: What it reveals and why it is important for formal linguistic theories. *Language and Linguistics Compass* 13(12). 1–19.

Mathesius, Vilém. 1947. *Čeština a obecný jazykozpyt [Czech language and general linguistics]*. Prague: Melantrich.

Miller, Jim & Regina Weinert. 1998. *Spontaneous spoken language: Syntax and discourse*. Oxford, UK: Oxford University Press.

Montrul, Silvina. 2010a. Dominant language transfer in Spanish L2 learners and heritage speakers. *Second Language Research* 26(3). 293–327.

Montrul, Silvina. 2010b. How similar are adult second language learners and Spanish heritage speakers? Spanish clitics and word order. *Applied Psycholinguistics* 31(1). 167–207.

Montrul, Silvina. 2016. *The acquisition of heritage languages*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139030502.

Napoli, Donna Jo & Rachel Sutton-Spence. 2014. Order of the major constituents in sign languages: Implications for all language. *Frontiers in Psychology* 5.

Newmayer, Frederick J. 2000. On the reconstruction of "proto-world" word order. In Chris Knight, Michael Studdert-Kennedy & James Hurford (eds.), *The evolutionary emergence of language*, 372–388. New York, NY: Cambridge University Press.

Polinsky, Maria. 2008. Heritage language narratives. In Donna M. Brinton, Olga Kagan & Susan Bauckus (eds.), *Heritage language education: A new field emerging*, 149–164. London: Routledge.

Polinsky, Maria. 2018. *Heritage languages and their speakers*. Cambridge, UK: Cambridge University Press. DOI: 10.1017/9781107252349.

Polinsky, Maria & Olga Kagan. 2007. Heritage languages: In the "wild" and in the classroom. *Language and Linguistics Compass* 1(5). 368–395.

Polinsky, Maria & Gregory Scontras. 2020. Understanding heritage languages. *Bilingualism: Language and Cognition* 23(1). 4–20. DOI: 10.1017/S1366728919000245.

Prada Pérez, Ana de & Diego Pascual y Cabo. 2012. Interface heritage speech across proficiencies: Unaccusativity, focus, and subject position in Spanish. In Kimberly Geeslin & Manuel Díaz-Campos (eds.), *Selected proceedings of the 14th Hispanic linguistics symposium*, 308–318. Somerville, MA: Cascadilla.

Ross, John R. 1967. *Constraints on variables in syntax*. Cambridge, MA: MIT. (Doctoral dissertation).

Sagarra, Nuria, Liliana Sánchez & Aurora Bel. 2019. Processing DOM in relative clauses: Salience and optionality in early and late bilinguals. *Linguistic Approaches to Bilingualism* 9(1). 120–160.

Sasse, Hans-Jürgen. 2001. Typological changes in language obsolescence. In Martin Haspelmath (ed.), *Language typology and language universals: An international handbook*, 1668–1677. Berlin, Germany: Walter de Gruyter.

Sekerina, I. 2003. Scrambling and processing: Dependencies, complexity and constraints. In S. Karimi (ed.), *Word order and scrambling*, 301–325. Oxford, UK: Blackwell.

Sekerina, Irina. 1997. *The syntax and processing of split scrambling constructions in Russian*. City University of New York. (Doctoral dissertation).

Sequeros-Valle, Jose, Bradley Hoot & Jennifer Cabrelli. 2020. *Judgment versus production data: clitic-doubled left dislocation in heritage spanish*, vol. 5. 47.

Sirotinina, Olga B. 1965. *Poryadok slov v russkom jazyke [word order in Russian]*. St. Petersburg, Russian: SGU.

Slioussar, Natalia. 2007. *Grammar and information structure: A study with reference to Russian*. Utrecht, Netherlands: LOT Publications.

Titov, Elena. 2020. Optionality of movement. *Syntax* 23(4). 347–374.

van Osch, Brechje & Petra Sleeman. 2018. Subject position in Spanish as a heritage language in the Netherlands: External and internal interface factors. In Anna Gavarró (ed.), *On the acquisition of the syntax of Romance*, 187–214. Amsterdam: John Benjamins.

Westergaard, Marit & Terje Lohndal. 2019. Verb second word order in Norwegian heritage language: Syntax and pragmatics. In David W. Lightfoot & Jonathan Havenhill (eds.), *Variable properties in language: Their nature and acquisition*, 91–102. Washington, DC: Georgetown University Press.

Yokoyama, Olga. 1986. *Discourse and word order*. Amsterdam: John Benjamins.

Zapata, Gabriela C., Liliana Sánchez & Almeida Jacqueline Toribio. 2005. Contact and contracting Spanish. *International Journal of Bilingualism* 9(3-4). 377–395.

Zuban, Yulia, Maria Martynova, Sabine Zerbian, Luka Szucsich & Natalia Gagarina. 2021. Word order in heritage Russian: Clause type and majority language matter. *Russian Linguistics* 45. 253–281. DOI: 10.1007/s11185-021-09246-1.

Chapter 5

Expanding structures while reducing mappings: Morphosyntactic complexity in agglutinating heritage languages

Terje Lohndal[a,b] & Michael T. Putnam[c,d]
[a]NTNU - Norwegian University of Science and Technology [b]UiT The Arctic University of Norway [c]Penn State University [d]University of Greenwich (CREL)

Research on heritage language grammars to date provides overwhelming support for the general stability of their syntactic systems, while the status of their morphology can vary considerably. In this chapter we offer remarks on the morphological complexity of agglutinating heritage languages, taking a closer look at a number of phenomena in Labrador Inuttitut, Cherokee, and American Hungarian. Four important findings emerge from our review: First, these phenomena align with previously documented and observed patterns in heritage language morphology (Polinsky 2018, Putnam et al. 2021). Second, heritage language morphology maintains a significant degree of complexity, even in languages found to be in a moribund state (Bousquette & Putnam 2020). Third, adopting an exoskeletal approach to morphosyntactic decomposition and complexity (Lohndal & Putnam 2021), we observe trends towards larger syntactic structures (for lexicalization), and inversely a reduction in the inventory of exponency. Fourth, we observe a general trend in the "shrinking" of computational domains for lexicalization and movement operations.

1 Introduction

Morphology proper has been – and continues to be – a hotly contested and embattled domain of linguistic inquiry. One particular debate in which morphology tends to rear its (ugly) head time and again centers on attempts to determine and measure *complexity* in linguistic systems. Of course, an unavoidable prerequisite

Terje Lohndal & Michael T. Putnam. 2024. Expanding structures while reducing mappings: Morphosyntactic complexity in agglutinating heritage languages. In Maria Polinsky & Michael T. Putnam (eds.), *Formal approaches to complexity in heritage languages*, 101–132. Berlin: Language Science Press. DOI: ??

to any of these arguments requires the establishment of two important points: (1) exactly *how*, i.e., according to what set of heuristics, one intends to measure complexity/simplicity, and (2) exactly *where* morphological processes and operations are located within a particular architecture, i.e., is morphology afforded a unique modular status or is it distributed across other modules? Speaking to the first point, arriving at an agreed upon metric to measure morphological complexity has proven to be quite elusive (Arkadiev & Gardani 2020, Stump 2017, Anderson 2015), with some linguists resorting to frequency and entropy conditions (Ackerman & Malouf 2013), while others turning to particular processes, such as *demorphologization*, as a sign of an increase (or decrease) in systemic morphological complexity (Mansfield & Nordlinger 2020, Hopper 1990). An outcome of this research that bears repeating is that complexity in detailed and accurate description extends beyond inflectional morphology and may reach all the way into syntax. The lack of a universally agreed upon complexity metric in connection with morphology spills over into theory-building efforts (Mithun 2020). In this chapter, we adopt the approach that the division between "morphology" and "syntax" is largely arbitrary (Haspelmath 2011), and further develop a model of morphological complexity in heritage language grammar initially sketched out by Lohndal & Putnam (2021).

An additional confound arises when we analyze the morphology of bi- and multilinguals, and heritage bilinguals in particular. Experimental research overwhelmingly confirms the integrated nature of the bilingual lexicon (Kroll & Gollan 2014, Putnam, Carlson, et al. 2018). The current consensus seems to advocate a general reduction in complexity in heritage morphology (Scontras et al. 2015, 2018), and some argue that these outcomes are an indication of different developmental trajectories in heritage language morphology (Berdicevskis & Semenuks 2020). In spite of these suggested trends towards "simplification" (again, however one determines to measure this out), it is difficult, and in some instances even erroneous, to insinuate that (some level) of complexity is not maintained in heritage language grammars. Evidence in support of this claim can be found in moribund heritage grammars (Bousquette & Putnam 2020) as well as receptive bilinguals (Sherkina-Lieber 2015). In this chapter we take a closer look at the concept of morphological complexity in a particular set of heritage languages, namely, those that can be typologically classified as being *agglutinating*. In the remainder of this paper, we classify *polysynthetic* languages as a sub-class of agglutinating ones.[1]

[1]Polysynthetic languages are by nature somewhat difficult to disambiguate from agglutinating ones (Baker 1996, Mattissen 2004), however, one key trait that these former languages exhibit in more detail than the latter is complex phonological alternations in addition to agglutination (again, among other properties).

Although it has been suggested that irrespective of their typological distinction, the morphology of heritage bilinguals seems to follow a limited number of observable, perhaps *universal* trends (Polinsky 2018, Polinsky & Scontras 2020, Putnam et al. 2021), there are two structural outcomes in particular that we would like to examine in more detail. The first of these involves the "teasing apart" of synthetic morphological forms in favor of more analytic forms. Consider the following data from the polysynthetic heritage language Caddo (from Chafe 1976: 74, cited by Melnar 2005: 5–6). The examples in (1) involve posture stative predicates, which are expressed as single, morphologically complex units (or, *words*).

(1) a. háh?**awis?**nássa?
IND-sitting-be.cold-IMPFV
'[he] is cold while sitting.'

b. háh?**ánkis?**nássa?
IND-standing-be.cold-IMPFV
'[he] is cold while standing.'

c. háh?**ini**.nássa?
IND-lying-be.cold-IMPFV
'[he] is cold while lying.'

Chafe (1976) reported that expressing the posture stative predicates *?awis-* 'sitting', *?anikis-* 'standing', and *?ini-* 'lying' can now only be expressed by periphrastic counterparts, with the verbal element combined with a following copula. These structures consist of two *words* rather than just one, see (2).

(2) a. háh?ánássa? háh?**áwsa?**
IND-be.cold-IMPFV IND-sitting-be
'[he] is cold while sitting.'

b. háh?ánássa? háh?**ánkisa?**
IND-be.cold-IMPFV IND-standing-be
'[he] is cold while standing.'

c. háh?ánássa? háh?**ín?a?**
IND-be.cold-IMPFV IND-lying-be
'[he] is cold while lying.'

The second structural property of synthetic heritage languages that we consider here are cases whereby morphological exponents can be used in a wider set of contexts than before. As such, these exponents have a more generalized meaning

and they have lost the connection to the previously specified meaning. For now, we refer to such elements as *generalized exponence*. As an example of this, consider noun stems in Upper Tanana derived by the suffix *ł*. Many of the nouns denoting tools and instruments contain this morpheme, whose instrumental meaning in the heritage language has become opaque, or lost in most instances (data from Lovick 2020: 148).

(3) Upper Tanana noun stems derived with *ł*

a. *teeł* 'mat'
b. *eeł* 'trap'
c. *tsiił* 'bridge, weir'

These morphological processes raise interesting and timely questions for linguists interesting in gaining a better understanding of the formal properties of complexity in heritage language morphology.[2] The first case involves cases whereby morphemes are "decoupled" from their original host, leading to *increased analyticity* (Polinsky 2018: 183). This state of affairs suggests that some degree of relativized economy of representations may also be present (Scontras et al. 2015, 2018, Perez-Cortes et al. 2019, Polinsky & Scontras 2020, Putnam et al. 2021). The second scenario reveals the opposite trend, namely, the maintenance of morphological forms that have become, or are well on their way to becoming, semantically bleached. Here we examine these trends in tandem in agglutinating heritage languages with consideration of how these contribute to a unified narrative on complexity in heritage grammars.

As already noted, in this chapter we home in on properties and trends observed in the morphology of *agglutinating* heritage languages, with a particular focus on verbal elements. Two of the three languages reviewed are also polysynthetic. In §2 we briefly review common properties of heritage morphology that transcend typological class (Polinsky 2018, Putnam et al. 2021). We introduce our formal conceptualization of complexity in §3, building on an initial treatment set forth by Lohndal & Putnam (2021). In §4 we sketch out an analysis of aspects of verbal morphology of heritage Inuttitut, Cherokee, and American Hungarian, focusing on the the rise of *increased analyticity* (§4.1) and *generalized exponence* (§4.2) and how their presence contributes to our treatment of complexity in heritage languages at the syntax-morphology interface. We provide the sketch of

[2]Note that we are here treating indigenous languages and immigrant languages on a par. In terms of formal properties, we believe that this can be justified, although this of course does not entail that the community dynamics of these contexts often differ substantially.

an analysis of these tendencies in §5, showing how an exoskeletal approach to the morphology-syntax interface delivers a unique and promising perspective on addressing the notion of complexity in heritage languages. Finally, we conclude this chapter in §6, pointing out fruitful areas of research into the notion of (morphological) complexity in heritage language grammars moving forward.

2 Structural tendencies in heritage morphology

As discussed in detailed by Polinsky (2018: Ch. 5) and Putnam et al. (2021: §2), the morphological systems of heritage languages share a number of structural tendencies. Space and time prevent us from providing a comprehensive overview of these common traits, however, we do wish to highlight both these general trends and additionally what they mean for establishing a formal heuristic of complexity. Morphological systems found in heritage languages develop a propensity for one-structure-to-one-meaning mappings irrespective of the typological system of morphology a particular language (predominantly) adheres to. Still, as we discuss in the remainder of this paper, there are interesting puzzles that heritage languages with agglutinating and polysynthetic morphological systems pose for attempts to formalize notions of representational economy and complexity.

Putnam et al. (2021) identify five primary patterns observed in heritage morphology irrespective of typological classification: (i) transparency and salience of forms and structures, (ii) overregularization and overmarking, (iii) preference for analytical forms, (iv) avoidance of ambiguity and underspecification, and (v) minimal domains. Let us now discuss each of these in turn based on the discussion in Putnam et al. (2021), which the reader should consult for examples and references.

Transparency of forms and structures refers to the mapping between an underlying feature and a given exponent. The most transparent case is the case whereby one feature refers to one exponent, making both readily identifiable. Therefore, forms and structures that are transparent are easier to detect in complex morphological paradigms. Typically, transparent forms will win out at the expense of less transparent forms. One illustration of this comes from grammatical gender, where heritage speakers often struggle to assign grammatical gender to non-transparent nouns. Another one involves the complete loss of a structural form, which has been found e.g., for the subjunctive.

Overregularization occurs when a speaker overuses highly transparent and regularized forms. For example, a particular morphological case form may be overextended and generalized so that, say, the nominative is also used with direct objects. When other case features are lost, the nominative form becomes

more accessible and therefore easier to overuse. However, it is not the case that what is robust in the input is always retained in heritage grammars. Subject-verb agreement is a good example, which often is reported to be vulnerable despite its prevalence in the input. Overmarking, on the other hand, is argued by Putnam et al. 2021 to be a special case of overregularization whereby a particular form is used in contexts where it normally would not be used. An example from Polinsky (2018) involving heritage English is the overuse of the weak past tense *-ed*-marker, producing forms such as *dresseded, walkeded* and *sanged, wented*. That is, the weak past tense form is overmarked.

Turning to the preference for analytical forms, this refers to the preference for analytical and periphrastic forms compared to synthetic ones. We saw an example of this already in the introduction. This may come from a drive to establish one-to-one mappings between form and meaning, which in turn will provide more transparent mappings (cf. the first pattern). Ultimately, this preference may be due to a bias in children to assume a one-to-one mapping between form and meaning when acquiring a language, cf. Slobin (1973) and van Hout (2008). If so, it is no surprise that such a bias is accentuated in heritage language settings (cf. Polinsky 2018).

A consequence of the preference for one-to-one mappings is the preference in heritage speakers to avoid structural ambiguities. That is, speakers avoid instances where one form may be associated with different meanings. For instance, in languages with dative case, the multiplicity of functions and mappings of dative case, restructuring is prone to occur. Assuming that scope is structurally determined, another illustration is that speakers tend to go for surface scope readings as opposed to inverse scope. Furthermore, heritage speakers tend to avoid underspecification in cases where two or more segments form a paradigmatic opposition. Instead, they often go for fully valued features, in which case oppositions act as a "vaccine" against underspecification. Generally, heritage speakers opt for one pattern or a fully specified feature structure.

The last pattern is what is typically referred to as *minimal domains*. This refers to a preference for minimal domains of computation. For instance, adjacent forms are preferred over dependencies that span a distance. In more formal terms, this suggests that the domains of computation are somehow "smaller" in heritage grammars. One way in which this occurs is for example by a reduction in the functional spine, either through heads/features merging or through a head/feature being eliminated.

With these five patterns in mind, we next turn to an attempt to decompose the notion of complexity and thereby put it on a more formal grounding.

3 Formalizing complexity

As Lohndal & Putnam (2021) demonstrate, complexity has been a much discussed topic in general and in the context of heritage grammars. The general intuition has been that heritage grammars tend to exhibit reduced complexity, although as Lohndal & Putnam also point out, complexity is a much too coarse grained notion to be of much use in understanding the nature of heritage grammars. Instead, it is vital to decompose *complexity* into familiar concepts and notions. In this section, we will summarize the approach developed in Lohndal & Putnam (2021), where three concepts are argued to be central in modeling complexity in a formal system.

(4) Criteria for establishing complexity:
   a. Number of syntactic features
   b. Number of functional projections
   c. Mapping from syntactic features to exponents

The criteria in (4) require a decompositional approach to the lexicon and an architecture which clearly distinguishes between syntax proper and morphophonology. A general label for such approaches is *exoskeletal approaches*, that is, approaches whereby syntactic structures determine both grammatical properties and "the ultimate fine-grained meanings of lexical items themselves" (Borer 2003: 33).[3] Unlike many traditional approaches whereby lexical items feed syntactic operations, syntactic structures are built from atomic features which are then associated with morphophonological realizations. The latter invokes a realizational approach to morphology in which morphosyntactic properties determine inflectional morphemes (see Stump (2001) for more on this). The syntax consists of atomic units, often referred to as roots and formal features, which mark syntactic and semantic properties. In turn, these map onto morphophonological realizations, called *exponents*. This mapping is subject to controversy, with proposals ranging from Distributed Morphology (Halle & Marantz 1993) to Nanosyntax (Starke 2009). We will briefly review each of these approaches.

Within Distributed Morphology, a Vocabulary Item denotes the mapping between abstract features and exponents. This is illustrated in (5) (adopted from Embick 2015: 9):

[3]We won't provide an in depth discussion of the relative merits of exoskeletal approaches over endoskeletal approaches. Readers are referred to, among many, Borer (2005a,b), Lohndal (2014, 2019), Wechsler (2021) for discussion.

(5) Vocabulary Item $\underset{\textit{synsem-features}}{[\alpha\beta\gamma]} \leftrightarrow \underset{\textit{phonological exponents}}{\underbrace{/X/}}$

In (5), the features are characterized as *synsem-features*, which means that they are syntactic-semantic features. These features have an interpretation. There may also be purely syntactic features of the kind that triggers movement to a dedicated position (such as the EPP feature of Chomsky (1982) or an Edge Feature as in Chomsky (2008)). For present purposes, we will limit our attention to synsem-features. Vocabulary Insertion is the operation that inserts phonological material into functional morphemes. Such sound/meaning connections can be quite complex, it is rarely the case that one functional feature is paired with one phonological exponent. Nevertheless, within DM, syntactic terminals are generally the targets of Vocabulary Insertion, meaning that, canonically, a syntactic head corresponds to an exponent.[4] As Svenonius (2016: 205) remarks, within DM a syntactic head thereby corresponds to a morpheme.

Nanosyntax takes a different point of departure, as it assumes that the lexicon consists of trees, which is to say that lexical items correspond to entire constituents. On this approach, a syntactic head (or terminal) is submorphemic, which is to say that many or most morphemes (and thereby exponents) will span several heads (Starke 2009). This intuition has been further developed in a span-based theory of how syntactic structures are mapped onto functional and lexical words (see, among others, Svenonius 2016). On this approach, a span is "a contiguous sequence of heads in a head-complement relation" (Svenonius 2016: 205). Thus, this architecture takes as its basis a non-transparent mapping between features and exponents.

Despite significant differences, both DM and Nanosyntax are committed to the existence of a distinction between the underlying features in a system, their subsequent values, and the actual exponents that are associated and matched with them. For present purposes, we would like to highlight this commonality, as a similar distinction has been invoked to analyze second language data and data involving language mixing in various kinds of multilingual speakers (Lardiere 1998, 2008, Prévost & White 2000, Alexiadou 2017, Alexiadou & Lohndal 2018,

[4]This does not mean that deviations from this canonicity are not important. In fact, the majority of research in this domain deals with such deviations, as Embick (2015: 25) points out: "One of the main topics in the theory of the morpheme concerns the possible departures from the one-to-one ideal, as the implementation of analyses that take these departures into account. [...] the theoretical imperative in this domain is to account for the attested departures from the ideal, while maintaining the most transparent (=strongest) theory of sound/meaning connections possible".

Riksem 2017, 2018, Lohndal et al. 2019, Putnam et al. 2019, Riksem et al. 2019, Putnam 2020). The approach in Lohndal & Putnam (2021) aims to develop this line of reasoning further to also address the question of how to predict possible outcomes in heritage grammars, cf. (Polinsky & Scontras 2020) and Putnam (2020). Lohndal & Putnam argue that there are four main possible outcomes, as illustrated in (6).

(6) Relative to a given baseline, a feature can
   a. be retained in the same hierarchical position
   b. shift its hierarchical position
   c. be lost
   d. be (internally) restructured
      i. loss of (some) features
      ii. reconfiguration of features

Lohndal & Putnam apply the four possible outcomes in (6) to grammatical gender in Spanish-English and Norwegian-English heritage situations and illustrate how they are attested in various scenarios. If all gender features are retained, there are no changes to the number of features. However, the functional sequence in which these features appear may or may not be identical to the relevant baseline. If both the features and the functional sequence are retained, there are no observable differences compared to the baseline. A feature can be lost, though, so that the grammar employs fewer gender distinctions. Either the feature simply disappears or it can be fused with another feature, which may result in a reduced functional sequence. Lastly but not least, the mapping from features to exponents may also undergo change. That is, the rules governing the syntax-morphology mapping may change, either so that there are (i) fewer rules, (ii) more rules, or (iii) that the rules remain but are altered in visible ways. Exactly how this materializes in heritage languages is something that Lohndal & Putnam do not really address.

In the present chapter, we would like to address the syntax-morphology interface by particularly focusing on the nature of exponence. A starting point is provided by Siddiqi (2009), who argues in favor of a principle that he labels *Minimize Exponence*. He defines it as in (7).

(7) *Minimize Exponence*:
The most economical derivation will be the one that maximally realizes all the formal features of the derivation with the fewest morphemes.

An example of this is the contrast between the acceptable *ate* and the unacceptable **eated*. In the former, the root $\sqrt{eat}$ and the feature [PAST] are realized as one morpheme. In the latter, the same root and feature would be realized as two morphemes: *eat* and *-ed*. Based on an assumption regarding the economical nature of derivations, the principle in (7) dictates that the derivation with the fewest morphemes converges. Thus, we get *ate* and not **eated*. Interestingly, children are known to produce examples such as **eated*. For that reason, Alexiadou (2021) and Hein et al. (2022) argue that they follow a different principle, namely *Maximize Exponence*. Alexiadou (2021) defines it as in (8).

(8) *Maximize Exponence*:
Realize semantic features in the building blocks of a complex unit by using *one exponent* for each feature.

The difference between the two varieties of exponence is crucially related to transparency. Alexiadou (2021) formalizes this as in (9).

(9) Assuming two semantic features *C1, C2*, may be realized either together by a single exponent *En* (*Fusion*) or by two exponents *E1, E2*, then *E1, E2* is *more transparent* than *En* (e.g., *make break* is more transparent than *break*)

Obviously, there are also cases in between the minimal and maximal alternatives, which is what is typically known as *Multiple Exponence*.[5] In this case, it is possible to get forms such as *ate-d* as realizations of $\sqrt{eat}$ and [PAST]. As argued by Alexiadou (2021) and Hein et al. (2022), children often start out by assuming Maximine Exponence and then they eventually converge on the adult system which adheres to Minimize Exponence. On the way, Multiple Exponence may appear as an intermediate stage, displaying what is often considered redundant commission errors in child speech. As emphasized by Alexiadou (2021), Multiple Exponence may surface in processes of language contact and language change. This will become relevant later on in this chapter when we will demonstrate that this is indeed the case.

In §2, we saw that there are general trends irrespective of the particular typological make up of a language. This raises the question of whether the approach outlined in this section can be extended to non-European languages. To address this, we will now look at agglutinating languages, two of which are also polysynthetic, to see whether and to what extent the approach developed in this section can be extended to such languages.

[5] Caballero & Harris (2012: 165) define this as "the occurrence of multiple realizations of a single morpho-semantic feature in a domain".

4 Trends in agglutinating heritage verbal morphology

In this section we zero in on heritage language verbal morphology with the intention of evaluating how these phenomena contribute to our discussion of complexity in heritage languages. We ground our empirical focus on data from three heritage speaker language communities, (i) Labrador Inuttitut, (ii) Cherokee, and (iii) American Hungarian. Labrador Inuttitut is largely a moribund heritage grammar, which has a high number of *receptive bilinguals*, i.e., those who are able to parse representations for comprehension (with varying degrees of success) but cannot speak the language anymore (without severe difficulty).[6] Cherokee is a Southern Iroquoian language related to Northern Iroquoian languages (e.g., Seneca, Oneida, and Mohawk). We acknowledge from the outset that it may be an extension to consider Cherokee a 'heritage language'; however, the conditions under which it is acquired and the sociolinguistic domains in which it is commonly used do share a number of relevant traits, which we elaborate on below. We round out our treatment of complexity in agglutinating languages by taking a closer look at morphological phenomena in three different American Hungarian communities, namely, (i) translation task data from three generations of heritage speakers in the San Francisco Bay area (Tóth 2007), (ii) naturalistic data from 2 generations of speakers in McKeesport, PA (Fenyvesi 2000), and (iii) spontaneous data (ca. 18 hours of recordings) from six Hungarian-English bilingual children (aged 7–9) studied in Bolonyai (2007). In spite of the fact that these sources span across generations of speakers of American Hungarian, the general empirical trends appear in both populations of speakers, providing further evidence of Putnam, Pascual y Cabo, et al.'s (2018) charge to treat these data sources on a par with one another.

This section is structured around the observable trends that we see across these three language communities: Labrador Inuttitut, Cherokee, and American Hungarian. In our view, these trends can be divided into two categories: Increased analyticity, which is the topic of §4.1, and what we call *generalized exponence*, which we treat in §4.2.

4.1 Increased analyticity

We start with American Hungarian spoken in the San Francisco Bay Area (Tóth 2007). Compared with other heritage language communities in the US, American Hungarian speakers in the Bay Area are a relatively young group, with the main

[6] See Sherkina-Lieber (2020) for a review of the literature on receptive bilinguals, including a classification of this group of bilinguals.

thrust of migration to California taking place after 1956. Most Hungarians left their homeland for politic reasons and continued to celebrate aspects of their ethnic heritage, such speaking Hungarian. As can be expected, over the course of multiple generations, a number of structural trends mark the language of 2nd and 3rd generation speakers, which serve as our empirical focus here. Twenty informants ($n = 20$) participated in a translation task of 1,000 sentences in Tóth's (2007) study, which he compared with European Hungarian as a "baseline."

One of the emerging structural trends he noted – especially in the translations produced by 3rd generation American Hungarian speakers – was a tendency towards producing more analytical morphological forms. In the examples in (10a) and (10b) below, the morphemes that mark number are accompanied by pronouns, resulting in redundancy that is not required in European Hungarian.

(10) Additional proforms, American Hungarian (Tóth 2007: 169)

a. *Én* nem tud-*om*, mért ő nem men-t-∅ haza.
I not know-1SG why he not went-PST-3SG home
'I don't know why he went home.'

b. *Ők* tud-t-*ák* hogy őneki rossz kedve volt.
they know-PST-3PL that him.DAT bad his.mood was
'They knew that he was in a bad mood.'

Although the above examples in (10) can be interpreted as leading to more analytic, and hence, transparent one-form to one-meaning mappings, they also result in what we above labeled multiple exponence. As a case in point, in (10a) the proform *Én* 'I' duplicates the same grammatical information as the suffix *-om* (1SG). The redundancy in exponence resembles the doubling of verbal elements with copulas in Caddo mentioned in the introduction (cf. (2); Chafe 1976, Melnar 2005). Note that multiple exponence can be compatible with transparency: a particular feature corresponds to a particular exponent, and in environments of multiple exponence, this happens twice in the same derivation.

Although increased analycity is the anticipated outcome in the development of heritage morphology, we sometimes observe the *opposite* trend.[7] For example, Bolonyai (2000) observes that whereas early system morphemes, i.e., those which do not assign or receive theta roles, remain relative stable in American Hungarian (when compared with European Hungarian as a baseline), late system morphemes, i.e., those that entail functional information, pose more difficulties

[7]We thank Oksana Laleko (p.c.) for bringing this matter to our attention.

in acquisition and can thus lead to a higher degree of divergent structures.[8] Preverbal elements in Hungarian have an ambiguous status: They are considered to be early system morphemes when they appear in a focus position (i.e., immediate proclitics of the verb they modify), but as late system morphemes when they must be positioned elsewhere, as as in the case of when certain auxiliary verbs are present. Intuitively, this definition of late system morphemes requires both syntactic and discourse information beyond the immediate verb phrase, thus requiring an extended domain of syntactic computation (cf. footnote 4).

As anticipated, the *late system* application of preverbs often lead to divergent structures (when compared with European Hungarian). Out of the 380 nontarget-like errors discussed in Bolonyai (2000), 225 (58.7%) are *late system morphemes*. The preverb *meg* 'me' appears directly before the auxilary/modal verb *tud-* 'can' in Standard Hungarian (11b), but in American Hungarian the preverb is cliticized directly onto the lexical verb *mond-* 'tell' as shown in (11a).

(11) Divergent preverb placement, American Hungarian (Bolonyai 2000: 96)

a. American Hungarian
Tud-om *meg*-mond-ani a mamá-d-nak
can-PRES-1SG-OBJ PREV-tell-INF the mom-POSS-2SG-DAT
'I can tell it to your mom.'

b. Standard Hungarian
Meg tud-om mond-ani a mamá-d-nak
PREV can-PRES-1SG-OBJ tell-INF the mom-POSS-2SG-DAT
'I can tell it to your mom.'

The challenge, of course, is to interpret what the cause of this apparent *increased syntheticy* might be. Although we fully acknowledge that data such as these require more rigorous treatment in the future, a distinct possibility motivating the restricted raising of the preverb *meg* in (11a) is that it is thematically marked by the predicate *mond-* 'tell', and remains structurally closer. If this postulation holds, this becomes a situation where some sort of *minimal domain*-preference wins out over increased analyticity; however, admittedly, this would require further investigation.

[8]This distinction is based on Myers-Scotton & Jake (2000) and their model of four types of morphemes. For them, early system morphemes are morphemes which are elected by content morphemes and together they form a semantic and structural unit. An early system morpheme is always realized inside of the maximal projection of the content morpheme that selects it. Late system morphemes, on the other hand, have as their main function to realize morphosyntactic information. As such, they provide the frame for the lexical-conceptual content which is specified in the lexicon.

4.2 *Generalized exponence*

Labrador Inuttitut is largely a moribund heritage grammar, which possesses a high number of receptive bilinguals. Sherkina-Lieber (2015) examines whether or not receptive bilinguals of Labrador Inuttitut are able to recognize and process semantic features such as tense, aspect, and agreement. Many actually displayed fluent-like comprehension of aspectual suffixes, subject-object agreement on verbs (in the form of suffixes), and past vs. future contrasts in tense suffixes. The data sample in (12) reviews the core tense morphology reflexes of this language (Sherkina-Lieber 2015: 36).

(12) a. Kaja-liu-juk
kayak-make-PART.3SG
'He is making a kayak.'

b. katat-tuk
fall-PART.3SG
'He (just) fell down.'

c. Kaja-liu-**laut**-tuk
kayak-make-**DPST**-PART.3SG
'He made a kayak (yesterday or earlier).'

These tense markings in (12) reflect the Eastern Canadian dialects of Inuktitut, which exhibit obligatory tense morphemes.[9] Present tense, which appears in (12a), is unmarked, and verbs without an overt tense morpheme are interpreted to take place during speech time. Achievement events, such as (12b), are an exception to this. By default, these predicates are perfective, and whenever they appear without a tense marker, they are interpreted as eventualities that took pace in the immediate past. To mark tenses other than the (unmarked) present, affixes are required. In example (12c) a tense morpheme indicating *distant past* appears immediately before agreement and mood inflection.

To test the quality of receptive bilinguals' representations of morphology in Labrador Inuttitut, twenty ($n = 20$) participants listened to 100 mini-stories read by a fluent native speaker of the language. These stories contained 84 target items, with 40 of them related directly to tense. The sentences were constructed in such a manner that if the informant was unable to process the target morpheme, the sentence would be ambiguous for them. Four tense morphemes were

[9]West Greenlandic Inuktitut, however, is assumed to not have obligatory tense morphemes, which are sometimes described in the literature as bound adverbials or aspectual markers.

directly targeted in this study: (i) distant past *-lauC-*, (ii) same day past *-kKau-*, (iii) same-day future *-niaC-*, and (iv) distant future *-lâC-*. The forced-choice comprehension questions had two options, depending on the targeted contrast. With respect to the contrast between same-day past vs. same-day future, informants encountered the question *Did X already V, or will X V soon?*, and to test the distinction between distant past vs. distant future they were forced to answer the following question, *Did X already V, or will X V later?*

As expected, proficiency in the heritage language, even in the case of receptive bilinguals, played a significant role. Whereas some of the the speakers with the lowest proficiency in Labrador Inuttitut failed make any distinctions with regard to the temporal morphology in the recordings, those with higher proficiency were able to successfully make the distinction between past, present, and future. Where all informants of this study struggled was in the ability use both TENSE and REMOTENESS in tandem to identify all four temporal morphemes tested in this study. We will return to a theoretical analysis of this in §5 below.

Turning to Cherokee, Peter et al. (2008) engage in an interesting study looking at children acquiring this language in the Cherokee Nation kindergarten immersion program (as documented on the Cherokee Kindergarten Immersion Language Assessment, C-KILA). Cherokee is classified as "severely endangered" by UNESCO. This means that it is a language mostly spoken by the grandparental generation and upward, and only by a minority of the population. Given this context, we believe it is warranted to treat Cherokee as a heritage language for these children.

Let us look at verbs in Cherokee. A minimum of two parts are required for all verbs in this language: (i) a *verb stem*, consisting of the a root and tense and aspect markers and (ii) a *pronominal prefix* (either Set A or B) that indicates the agent of the event.[10] Here are three examples from Peter et al. (2008: 174–175) of 3rd person, continuous present verbs in Cherokee.

(13) a. ga-thiha
3A.SG-sleep:PRC
'S/he is sleeping.'

b. ani-aditasga
3A.PL-drink:PRC
'They are drinking.'

[10]See Montgomery-Anderson (2015: Ch.3 & Ch. 4) for a detailed overview.

c. de-ga-hnogi′a
DST-3A.SG-sing:PRC
'She is singing.'

In all three examples above in (13), the verbal roots, *-tliha* 'sleep', *-aditasga* 'drink', and *-hnogi′a* 'sing', are augmented with additional morphological markers. All of these examples appear with a Set A pronominal prefix, which can be singular or plural (cf. 13a and 13b).[11] Note that the plural marker differs between Set A *ani-* and Set B *uni-*, and they are further subject to phonological integration processes depending on the root they merge with. To mark that these events are currently taking place, i.e., are in progress, a progressive morpheme has to be present (marked as PRC). As can be expected, if any of these elements are missing, these structures would be regarded as ill-formed or simply "incorrect".

Peter et al. (2008) describe the production of thirteen ($n = 13$) children in order to assess whether or not these children are able to use third person singular and plural present tense continuous verbs – such as those introduced in (13) immediately above. Peter et al. (2008: 175) focused on the following three aspects of the children's production of present continuous verb morphology: (i) the appropriateness of the verb root in relation to the picture shown in the task, (ii) the accuracy of the verb stem in combination with present tense and continuous aspect, and (iii) the accuracy of third person singular and plural pronominal prefixes, and the distributed pre-pronominal prefix when warranted.

Although these children had begun to apply verbal morphology in their production, there were noticeable "common errors" that surfaced in these aggregate data when compared with baseline forms provided by the eldest generation of Cherokee speakers. Aside from the expected errors that would result from children not having yet acquired the appropriate verb (and hence, lacking the verbal root in their lexicon), there were two sets of errors that emerged from the children's production: (i) inaccurate usage of pronominal and pre-pronominal prefixes, and (ii) inaccurate usage of tense and aspect markers. In the remainder of our discuss of the data from the children acquiring Cherokee, we focus on the former of these categories.

Breaking things down a bit further, there are three subclassifications of "divergent" structures – again, when compared with baseline forms provided by the eldest generation of Cherokee speakers – highlighted by Peter et al. (2008) in connection with inaccuracies involving pronominal and pre-pronominal prefixes.

[11] In (13c) we find an additional prefix that precedes the Set A pronominal prefix. The pre-pronominal element is a distributive (DST) marker, since the act of singing in Cherokee is a distributive event.

The first of them concerns overgeneralizations of the 3rd person plural pronominal prefix *ani-* to cover *all* plural cases. Applying *ani-* in all plural contexts ignores the set to which a verb belongs, and whether or not a verb requires a distributed pre-pronominal prefix. Second, Peter et al. (2008) observed the overuse of the 3rd person singular pronominal prefix *ga-*: "Of the 104 obligatory occasions that children had to produce a plural pronoun marker, they used the singular prefix 11 times, or in 11% of the unit counts" (Peter et al. 2008: 178). Third, the children were somewhat inaccurate with respect to their use of the distributed pre-pronominal prefix, although given the relatively low number of verbal tokens ($n = 7$) on the C-KILA materials that require the distributive marker, the thrust of this claim requires further investigation.

Shifting to research on the verbal complex of American Hungarian in McKeesport, PA, Fenyvesi (2000) analyzes naturalistic production of 20 speakers in the area, with 4 of them being 1st generation immigrants, and the other 16 2nd generation heritage speakers (predominantly monolingual speakers of American English). One feature that stands out in these data relates to European Hungarian's complex definiteness system. Definiteness marking involves the integration of separate person, number, and definiteness markers. As a result, the use of a prenominal definite determiner may not license the definite verbal suffix (see 14a). Intransitive verbs do not receive definite marking. According to Fenyvesi (2000), this element of the grammar has been simplified in McKeesport, PA American Hungarian. In this variety of American Hungarian, the determiner and the verbal suffix that indicates definiteness are congruent with one another (see 14b).[12]

(14) "Simplified" definite marking, American Hungarian (Fenyvesi 2000: 97)

a. European Hungarian
az öreg-ek meg-hal-t-ak
the old-PL PVB-**die-PST-3PL.INDEF**
'the old people died'

b. American Hungarian
az öreg-ek meg-hal-t-ák
the old-PL PVB-**die-PST-3PL.DEF**
'the old people died'

This finding is not unique to the McKeesport, PA American Hungarian speakers' grammars; Tóth (2007) finds additional evidence of similar phenomena in

[12]Hungarian <a> = /ɒ/ and <á> = /a:/.

his data, e.g., the extension (*overmarking*) of the *ik*-morpheme to verbs not originally included in this class (§4.6.14) (15), missing verb conjugations (§4.6.10) (16), and incorrect and inconsistent case marking on nouns (§4.6.1.1) (17). The examples of overmarking of *ik*-verbs in (15) both involve the verb to cook (*főz*). As for the missing verb conjugations, in (16a), the 1SG future auxiliary *-ok* is missing from *fog* 'will' in the conditional clause, and the 2SG past indicative suffix is missing from *kérdezt-* 'asked' in (16b). Tóth's (2007) findings on inconsistency case marking in SF Bay Area Heritage Hungarian are extensive, warranting a separate detailed analysis in their own right. In example (17a), the absence of the dative case marking suffix *-nek* renders the noun *nő* 'woman' in the nominative. The example in (17b) is unique, but crucially relevant for our discussion here. According to Tóth (2007: 140), this informant produced all three case endings, i.e., INSTRUMENTAL: *-mal*, ACCUSATIVE: *-at*, and DELATIVE: *-ról*, in succession in an attempt to express something equivalent to English *thinking about*. We interpret this as a canonical instances of the Maximize Exponency axiom previously introduced in this chapter.

(15) *Overmarking* of *ik*-verbs

a. Af férfi, aki főz-*ik*, az az én férjem.
he man who is-cooking that the my husband
'The man who is cooking is my husband.'

b. Az ember, aki főz-*ik*, a férjem.
the man who is-cooking the my-husband
'The man who is cooking is my husband.'

(16) Missing verb conjugations

a. Ez egy ember fog-*ok* jönni, én fog adni neki pénzet.
this one man he-will come.INF I will give.INF him.DAT money.ACC
'If this one man will come, I will give him money.'

b. Mit kérdezt-*(t)él* a rendőr?
what.ACC asked the policeman
'What did the policeman ask?'

(17) Inconsistent case mismatches

a. Majd jön egy alkalom, mikor meg fogja köszönni a
later comes an occasion where PVB she-will.DEF thank.INF the
nő-*nek*, aki szembe lakik vele.
woman.NOM who acrss she-lives she.INSTR
'Later an occasion will come, when she will thank the woman she lives across from.'

b. Vándoroltam a szabad téren, gondoskodtam a
I-hiked the open space.SUPERESS I-cared-for the
szomszédom-**mal/-at/-ról**.
my-neighbor.-INSTR/ACC/DELAT
'As I hiked the open space, I thought about the neighbor I cared about.'

As we highlight below, these observable patterns are also present in the production of American Hungarian-speaking children.

The final piece of empirical evidence in the American Hungarian verbal complex concerns instances of language mixing, in which an English-based $\sqrt{\text{root}}$ combines with Hungarian morphosyntax. Based on previous published research and her extensive analysis of an individual child, Bolonyai (2005) highlights the requirement of an additional verbalizing derivational suffix *-l* (*-ol, -el, -öl*) (abbreviated: VBZ) that appears in combination with English-origin verbs, as illustrated in (18):

(18) Obligatory verbalizers in bilingual Hungarian (Bolonyai 2005: 317)

a. *cover*-**ol**-ja
cover-**VBZ**-PRES.3SG.DEF
'[it] covers [it]'

b. el-*explain*-**el**-ni
PV-explain-**VBZ**-INF
'to explain'

c. fel-*pick*-**ol**-t-am
PV/up-pick-**VBZ**.PAST.1SG
'I picked him up.'

d. fel-*réz*-**ol**-t-am
PV/up-raise-**VBZ**.PAST.1SG
'[I] raised up.'

One of the key questions pursued by Bolonyai (2005) centered on whether or not English-Hungarian bilingual children would also require these the obligatory presence of the derivational suffix *-l* in "mixing contexts," assuming that children also would mix English and Hungarian as shown in (18). Approximately 30 hours of longitudinal naturalistic bilingual conversations of an individual English-Hungarian bilingual child recorded across a span of 7 years from age

3;7 to 11;3 served as the empirical base of this investigation. The recordings consisted of primary stages: *Stage One* (20 hours) included recordings from ages 3;7–5;10, while *Stage Two* (10 hours) focused on ages of 6;8–11;3. Two findings are of particular relevance for us: First, during Stage One the vast majority of English-origin roots did not occur with the expected Hungarian verbalizing suffix (86/92; 93.5%). During Stage Two, we observe the inverse of this behavior, with the expected verbalizer occurring in 95.5% of possible forms (105/110). Second, during Stage Two the number of possible candidate forms doubled. These findings suggest that the acquisition of this language-mixing behavior observed in bilingual Hungarians is not acquired (fully) before the age of at least 6;0, which indicates this as a potential vulnerable domain of grammar in a heritage language setting where language mixing occurs.

Summarizing, what we have seen in this section is that features are not necessarily lost in heritage speakers; rather, they can be "thinned out" or generalized across contexts. The latter idea refers to the fact that distinctions can be lost, when e.g., the same form is used to expone features that traditionally have separate forms. For instance, the same exponent may express both [PRESENT] and [PAST]. However, an alternative scenario is that speakers eliminate the distinction between [PRESENT] and [PAST] and just have one feature, say [TENSE], underlyingly. As Riksem (2017) argues at length, it is typically difficult to find convincing arguments in favor of either story given that the outcome in terms of exponents is identical. Lastly, when distinctions are lost, that can in turn lead to a re-organization in the feature inventory itself for a particular property. In the next section, we will provide a theoretical grounding of these general findings.

5 Decomposing complexity in agglutinating heritage languages

In the previous section, we saw that agglutinating and also polysynthetic heritage languages display a tendency towards increased analyticity (i.e., they tend to isolate more), and that they more often resort to what we have referred to as generalized exponence: They use one exponent to cover more contexts compared to a given baseline. Let us review these two generalizations and consider their theoretical implications. In particular, we will propose some implementations in terms of Distributed Morphology to illustrate the mappings between syntactic features and morphophonological exponents.

At a general level, it seems like the situation with Labrador Inuttitut and Cherokee are more or less similar in their reduction of their feature inventory. The

former involves a loss of a feature. Recall that Sherkina-Lieber (2015) compares heritage Labrador Inuttitut towards a baseline that she labels "full" Labrador Inuttitut. In terms of (semantic) features of the tense morphemes, she proposes the following figures which detail the differences between the two varieties. Figure 1 represents the (Sherkina-Lieber 2015: 44) the "full" variant, which exhibits feature distinctions for Past, Present, Future, as well as remote temporal distinctions.

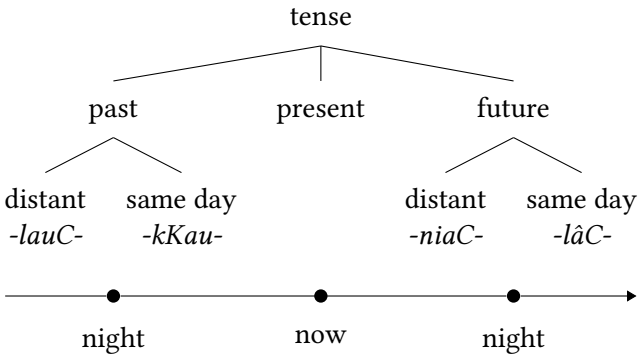


Figure 1: Semantic tense features in "full" Labrador Inuttitut

Contrast this situation with the feature inventory presents in "heritage" Labrador Inuttitut (Sherkina-Lieber 2015: 44) in Figure 2.

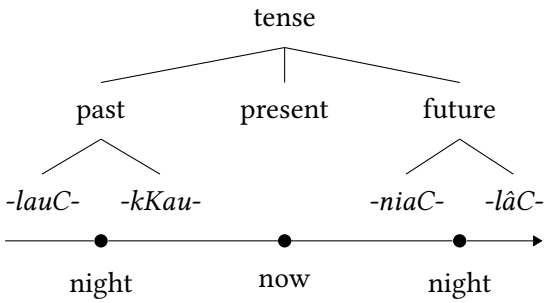


Figure 2: Semantic tense features in "heritage" Labrador Inuttitut

Comparing the feature inventories of the "full" and "heritage" variants with one another, we see that each tense morpheme is specified for both time and remoteness, whereas in the heritage speakers, each tense morpheme only has a specification for time. That is, no information about remoteness is available.

That means that even if these speakers have two markers for each tense available, they are not able to distinguish them. As Sherkina-Lieber (2015: 45) points out, remoteness is likely to be vulnerable due to a lack of any analogue in English. Features of a weaker language which have no counterpart in the dominant language tend to be more vulnerable across heritage languages. Here we also see generalized exponence at work: There are two morphemes available that are no longer associated with an underlying feature distinction. In that sense, the morphemes contain fewer features, even though the number of exponents remain the same.

As for Cherokee, we observe the application of the 3rd plural prenominal prefixes in all environments, which represents a prototypical case of *overgeneralization*. That is, speakers are no longer sensitive to verbs taking different prefixes. Again we see a case of generalized exponence where speakers are no longer sensitive to a distinction that exists in the baseline. That is, the distinction between Vocabulary Insertion rules in (19) is eliminated and speakers instead adopt the rule in (20). For expository convenience, we have just utilized a binary feature [±A] to refer to the distinction between Set A and Set B.

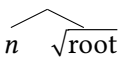
(19) a. [NUMBER:PLURAL, PERSON: 3, +A] ⟺ *ani-*
b. [NUMBER:PLURAL, PERSON: 3, –A] ⟺ *uni-*

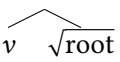
(20) [NUMBER:PLURAL, PERSON: 3] ⟺ *ani-*

In addition, as described above, some speakers also use the singular prefix as opposed to the plural, possibly indicating the beginning of a merger of the singular and plural pronominal prefixes.

Turning to the definiteness system in American Hungarian, we see that definiteness marking in this particular heritage language appears to have been "simplified". This difference, when compared to the baseline of European Hungarian, can be viewed as yet another instance of *generalized exponence*, in the sense that speakers try to unify agreement within a minimal domain (viz. the nominal phrase in this case).

American Hungarian also lacks obligatory verbalizers in the speech of child bilinguals, or at the very least, are mastered only after the age of 6. Theoretically, this is an interesting case, especially from the perspective of Distributed Morphology. Recall, that roots exist in the lexicon as a-categorial, feature-less entities that are receive categorial status by virtue of merging with a categorizer in the syntax. An illustration of this is provided in examples (21–22) for nouns and verbs, respectively.

(21)


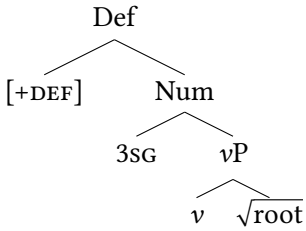
(22)


In many languages, such categorizers have overt exponents. This can easily be seen in English, where morphemes such as *-en, -age, -al* realize the categorizers *v, n* and *a* for words such as *darke, marriage, global*. In §4.2, we saw that American Hungarian shows the tendency of not retaining these obligatory verbalizers as anticipated in instances of language-mixing. By assumption, such categorizing heads are present early on in acquisition. This state of affairs means that for some reason American Hungarian speakers are tolerating what is effectively a "silent" head for some time, at least throughout some stages of development. This behavior at first glance appears to contradict an established claim in the heritage language literature arguing that heritage speakers struggle with silence, i.e., elements in syntactic structure that do not map to a corresponding phonological exponent, in general (see Polinsky (2018) for an extensive review), but this difference highlights the importance of also considering typologically diverse languages when crafting generalizations.

To illustrate how we would formalize this situation, consider example (18a), repeated below as (23):

(23) *cover*-**ol**-ja
cover-**VBZ**-PRES.3SG.DEF
'[it] covers [it]'

The verbalizing exponent *-ol-* in (23) is the morphophonological manifestation of the categorizing head *v*. The tree structure in (24) immediately below holds for American Hungarian expressions which contain and omit this verbalizing exponent.

(24)


While the features [+DEF] and 3SG undergo fusion and are spelled out as *ja* in (23), the Vocabulary Items for *v* are optional, at least in a developmental sense:

(25) a. $[v] \Longleftrightarrow ol$ [Overt realization of v]
b. $[v] \Longleftrightarrow \emptyset$ [Non-realization of v]

Another puzzle that emerges in connection with American Hungarian concerns the inclusion of additional proforms. On the one hand, we are dealing with one-to-one mappings, which we have argued are simpler than one-to-many mappings. However, this also leads to (i) redundancy, and (ii) the realization of smaller units (proforms) as "independent units" that used to be realized as suffixes. Classifying the occurrence of the realization of these (additional) proforms in terms of either reduced or increased complexity is quite difficult, but based on the decomposition of syntactic structures and their morphological reflexes proposed in §3, we adopt the position that these mappings result in more straightforward instances of correspondence even if they ultimately lead to redundancy and additional independent units. This may be an instance of what Alexiadou et al. (2021) calls MAXIMIZE EXPONENCE (8), which holds that for each semantic feature at least one exponent should be realized. Their evidence come from a very different empirical domain, and with converging evidence from multiple speaker groups this suggests that we are dealing with a possibly deep generalization regarding the interface between features and exponents.

Lastly, we provide a few remarks on instances of divergent preverb placement in American Hungarian. Recall from our discussion of the data in (11) that whereas the the preverbal particle *meg* raises to precede the auxiliary in European Hungarian, this fails to leave the verbal phrase in American Hungarian, appearing as a proclitic to the lexical verb. This situation illustrates the potential reduction in the movement of this preverbal particle, which can further be generalized to be a reduction in the computational domain of movement, i.e., a "minimal domain" effect. The latter has been demonstrated also in other cases, for instance long-distance binding (Gürel 2007, Kim et al. 2009, Putnam & Arnbjörnsdóttir 2015, Montrul 2016) and A-bar dependencies (Hopp et al. 2019). Further research should examine to what extend these "minimal domain" effects can be equated with some version of phase theory (see, e.g., Chomsky 2000, 2001, Gallego 2010), which is a fruitful topic that we leave for future inquiry.

Let us return to the typology proposed by Lohndal & Putnam (2021) in connection with an exoskeletal approach to complexity at the morphology-syntax interface, repeated in (26) for expository convenience.

(26) Relative to a given baseline, a feature can
a. be retained in the same hierarchical position
b. shift its hierarchical position

c. be lost
d. be (internally) restructured
   i. loss of (some) features
   ii. reconfiguration of features

The data reviewed in §4 provide evidence for (26d) and illustrate how some features and feature distinctions can be lost or eliminated. In order to better understand the mapping from features to exponents, we have focused on what we have called *generalized exponence*, which illustrates a situation where heritage speakers use an exponent in more contexts than baseline speakers. It should be added that generalized exponence can often be found in situations of language change more generally, although space prevents us from addressing that issue here.

Zooming out, the data and analyses in this chapter suggest that the particular morphological type is not a decisive factor when it comes to decomposing complexity in heritage languages. The same basic mechanisms appear to be at work across morphological systems. As such, this aligns with the generalizations and conclusions arrived at in Putnam et al. (2021); namely, that irrespective of the morphological typology of a given language, or dyad of languages in the case of heritage speakers, similar patterns of morphological outputs in heritage languages will manifest themselves along the lines of the five tendencies outlined in §2. Theoretically, these data adduce further support for a late-insertion account of morphology whereby the syntactic features are independent of morphophonological exponents. The review of different approaches to the syntax-morphology interface in §3 demonstrates that even within late insertion approaches there are substantial differences. The data in the present chapter can easily be analyzed within Distributed Morphology.[13]

6 Conclusion

Our main purpose in this chapter sought to unite current theoretical analyses of heritage language morphology with discussions centered on notions of *complexity* in heritage languages (and linguistic systems more generally). For the sake of space and time, we zeroed in on two empirical phenomena found in agglutinating languages, namely, *increased analyticity* and *generalized exponency*.

[13]They are also compatible with a Nanosyntactic account, although as far as we can tell, the data do not appear to necessitate such an approach.

Unsurprisingly, these heritage languages show an increased amount of one-to-one mappings (*increased analyticity*) and a reduction in feature-exponent mappings (*generalized exponency*), as explicated in detailed reviews by Polinsky (2018: Ch.5) and Putnam et al. (2021). Interpreting these findings through the lens of an exoskeletal approach to the morphology-syntax interface as laid out by Lohndal & Putnam (2021) (see §3), three primary outcomes emerge: First, the relative size of syntactic objects appear to be somewhat larger in heritage language morphosyntax. That is, the units that are lexicalized are larger because of changes in the feature structure. Under this notion, the semantic content of some features that were associated with morphophonological exponency is *weakened*, leading to the eventual possible loss of these connections. Therefore, whereas the syntactic objects, i.e., the domains of syntactic structure that are lexicalized, may grow in size, the inventory of exponents they may be associated may be (significantly) reduced. In some respects, we witness a "hollowing out" of syntactic structure, which forces us to revisit exactly what "representational economy" (Scontras et al. 2015, 2018, Polinsky & Scontras 2020) means in the context of the morphology-syntax interface in heritage languages. Second, instances of *overmarking* and *redundancy* sometimes accompany the observable trend toward analyticity. Even if such structures represent a transient state of the heritage language grammar, they represent instances of *distributed exponency*, which reinforce that not all morphological developments in heritage languages can be interpreted as internally-motivated simplification strategies (Bousquette & Putnam 2020). Third, the reduction in computational domains with respect to potential spell-out domains plays a role in the appearance of this aforementioned distributed exponency as well as the obfuscation of preverbal marker placement in languages such as American Hungarian.

The next frontier would be to see how the ideas in this chapter extend to other agglutinating and polysynthetic languages, in particular the latter. Although studies on "obsolescence" in polysynthetic languages appear to exhibit similar morphological patterns (Mithun 1989, Gruzdeva & Vakhtin 2017), the interplay of complex phonological alternations adds yet an additional important, yet intriguing domain to the puzzle of systemic complexity in heritage languages.

Abbreviations

A	Set A	DELAT	delative
DAT	dative	DPST	distant past
DEF	definite	DST	distributive

IMPFV	imperfective
IND	indicative
INDEF	indefinite
INF	infinitive
OBJ	object (non-agentive)
PART	partitive
PL	plural
POSS	possessive
PRC	progressive
PRES	present tense
PST	past tense
PVB	preverbal element
SG	singular
SUPRESS	supressive
VBZ	verbalizer

Acknowledgments

Thanks are in order to Artemis Alexiadou for reading through and commenting on a final version of this chapter. We also owe a debt of gratitude to an anonymous reviewer for their challenging and probing questions and criticisms which ultimately strengthened this chapter. A special thanks to David Natvig and Ashley Pahis for some LaTeX troubleshooting.

References

Ackerman, Farrell & Robert Malouf. 2013. Morphological organization: The Low Conditional Entropy Conjecture. *Language* 89(3). 429–464.

Alexiadou, Artemis. 2017. Building verbs in language mixing varieties. *Zeitschrift für Sprachwissenschaft* 36. 165–192.

Alexiadou, Artemis. 2021. Multiple exponence: Sometimes two is better than one. *Talk given at 46. Österreichische Linguistik-Tagung, December 9-12.*

Alexiadou, Artemis, Imke Driemel, Johannes Hein, Fabienne Martin & Yining Nie. 2021. Formalization of commission errors: Underassociation and minimize exponence. *Ms., HU Berlin.*

Alexiadou, Artemis & Terje Lohndal. 2018. Units of language mixing: A cross-linguistic perspective. *Frontiers in Psychology* 9. 1719.

Anderson, Stephen R. 2015. Dimensions of morphological complexity. In Matthew Baerman, Dunstan Brown & Greville G. Corbett (eds.), *Understanding and measuring morphological complexity*, 11–26. Oxford: Oxford University Press.

Arkadiev, Peter & Francesco Gardani. 2020. Introduction: Complexities in morphology. In Peter Arkadiev & Francesco Gardani (eds.), *The complexities of morphology*, 1–21. Oxford: Oxford University Press.

Baker, Mark. 1996. *The polysynthesis parameter*. Oxford: Oxford University Press.

Berdicevskis, Aleksandrs & Arturs Semenuks. 2020. Different trajectories of morphological overspecification and irregularity under imperfect language learning. In Peter Arkadiev & Francesco Gardani (eds.), *The complexities of morphology*, 283–305. Oxford: Oxford University Press.

Bolonyai, Agnes. 2000. "Elective affinities": Language contact in the abstract lexicon and its structural consequences. *International Journal of Bilingualism* 4(1). 81–106.

Bolonyai, Agnes. 2005. English verbs in Hungarian/English codeswitching. In James Cohen, Kara T. McAlister, Kellie Rolstad & Jeff MacSwan (eds.), *ISB4: Proceedings of the 4th International Symposium on Bilingualism*, 317–327. Somerville, MA: Cascadilla Press.

Bolonyai, Agnes. 2007. (In) vulnerable agreement in incomplete bilingual L1 learners. *International Journal of Bilingualism* 11(1). 3–23.

Borer, Hagit. 2003. Exo-skeletal vs. Endo-skeletal explanations: Syntactic projections and the lexicon. In John Moore & Maria Polinsky (eds.), *The nature of explanation in linguistic theory*, 31–67. Stanford, CA: CSLI Publications.

Borer, Hagit. 2005a. *Structuring Sense*, vol. 1: In name only. Oxford: Oxford University Press.

Borer, Hagit. 2005b. *Structuring Sense*, vol. 2: The normal course of events. Oxford: Oxford University Press.

Bousquette, Joshua & Michael Putnam. 2020. Reassessing language death: Evidence from moribund grammars. *Language Learning* 70(S1). 188–225.

Caballero, Gabriela & Alice C. Harris. 2012. A working typology of multiple exponence. In Ferenc Kiefer, Mária Ladányi & Péter Siptár (eds.), *Current issues in morphological theory: (Ir)regularity, analogy and frequency. Selected papers from the 14th international morphology meeting, Budapest, 13–16 May 2010*, 163–188. Amsterdam: John Benjamins.

Chafe, Wallace L. 1976. *The Caddoan, Iroquoian, and Siouan languages*. Berlin/The Hague: Mouton.

Chomsky, Noam. 1982. *Some concepts and consequences of the theory of government and binding*. Cambridge: MIT Press.

Chomsky, Noam. 2000. Minimalist Inquiries: The Framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by Phase. In Michael Kenstowicz (ed.), *Ken hale: A Life in Language*, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-roger vergnaud*, 133–166. Cambridge, MA: MIT Press.

Embick, David. 2015. *The morpheme: A theoretical introduction*. Berlin: Mouton de Gruyter.

Fenyvesi, Anna. 2000. The affectedness of the verbal complex in American Hungarian. In *Language contact and verb complex of Dutch and Hungarian*, 94–107. Szeged: University of Szeged.

Gallego, Ángel J. 2010. *Phase Theory*. Amsterdam: John Benjamins.

Gruzdeva, Ekaterina & Nikolai Vakhtin. 2017. Language obsolescence in polysynthetic languages. In Michael Fortescue, Marianne Mithun & Nicholas Evans (eds.), *The Oxford Handbook of Polysynthesis*, 428–448. Oxford: Oxford University Press.

Gürel, Ayse. 2007. (psycho)linguistic determinants of L1 attrition. In Barbara Köpke, Monika S. Schmid, Merel Keijzer & Susan Dostert (eds.), *Language attrition: Theoretical perspectives*, 99–119. Amsterdam: John Benjamins.

Halle, Morris & Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In Ken Hale & Samuel J. Keyser (eds.), *The view from Building 20: Essays in honor of Sylvain Bromberger*, 11–76. Cambridge, MA: MIT Press.

Haspelmath, Martin. 2011. The indeterminancy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45(1). 31–80.

Hein, Johannes, Fabrienne Martin, Yining Nie & Artemis Alexiadou. 2022. Minimizing and maximizing exponence in child language. *Talk given at WCCFL 40, Stanford University, May 13-15*.

Hopp, Holger, Michael T. Putnam & Nora Vosburg. 2019. Derivational complexity vs. Transfer effects. Long-distance *wh*-movement in heritage and L2 grammars. *Linguistic Approaches to Bilingualism* 9(3). 341–375.

Hopper, Paul. 1990. Where do words come from? In William Croft, Keith Denning & Suzanne Kemmer (eds.), *Studies in typology and diachrony: Papers presented to Joseph H. Greenberg on his 75th birthday*, 151–160. Amsterdam: John Benjamins.

Kim, Ji-Hye, Silvina Montrul & James Yoon. 2009. Binding interpretation of anaphors in Korean heritage speakers. *Language Acquisition* 16. 3–35.

Kroll, Judith F. & Tamar H. Gollan. 2014. Speech planning in two languages: What bilinguals tell us about language production. In Matthew Goldrick, Victor Ferreira & Michele Miozzo (eds.), *The Oxford handbook of language production*, 165–181. Oxford: Oxford University Press.

Lardiere, Diane. 1998. Dissociating syntax from morphology in a divergent L2 end-state grammar. *Second Language Research* 14(4). 359–375.

Lardiere, Diane. 2008. Feature-assembly in second language acquisition. In Juana Liceras, Helmut Zobl & Helen Goodluck (eds.), *The role of formal features in second language acquisition*, 107–140. Mahwah, NJ: Lawrence Erlbaum.

Lohndal, Terje. 2014. *Phrase structure and argument structure: A case study of the syntax-semantics interface*. Oxford: Oxford University Press.

Lohndal, Terje. 2019. Neodavidsonianism in semantics and syntax. In Robert Truswell (ed.), *The Oxford handbook of event structure*, 287–313. Oxford: Oxford University Press.

Lohndal, Terje & Michael T. Putnam. 2021. The tale of two lexicons: Decomposing complexity across a distributed lexicon. *Heritage Language Journal* 18(2). 1–29.

Lohndal, Terje, Jason Rothman, Tanja Kupisch & Marit Westergaard. 2019. Heritage language acquisition: What it reveals and why it is important for formal linguistic theories. *Language and Linguistics Compass* 13(12). 1–19.

Lovick, Olga. 2020. *A grammar of Upper Tanana*, vol. 1: Phonology, lexical classes, morphology. Lincoln, NE: University of Nebraska Press.

Mansfield, John & Rachel Nordlinger. 2020. Demorphologization and deeping complexity in Murrinhpatha. In Peter Arkadiev & Francesco Gardani (eds.), *The complexities of morphology*, 52–80. Oxford: Oxford University Press.

Mattissen, Johanna. 2004. A structural typology of polysynthesis. *Word* 55(2). 189–216.

Melnar, Lynette R. 2005. *Caddo verb morphology*. Lincoln, NE: University of Nebraska Press.

Mithun, Marianne. 1989. The incipient obsolescence of polysynthesis: Cayuga in Ontario and Oklahoma. In Nancy C. Dorian (ed.), *Investigating Obsolescence: Studies in Language Contraction and Death*, 243–257. Cambridge University Press.

Mithun, Marianne. 2020. Where is morphological complexity? In Peter Arkadiev & Francesco Gardani (eds.), *The complexities of morphology*, 306–330. Oxford: Oxford University Press.

Montgomery-Anderson, Brad. 2015. *Cherokee reference grammar*. Norman, OK: University of Oklahoma Press.

Montrul, Silvina. 2016. *The acquisition of heritage languages*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139030502.

Myers-Scotton, Carol & Janice L. Jake. 2000. Explaining aspects of codeswitching and their implications. In Janet Nicol (ed.), *One mind, two languages: Bilingual language processing*, 91–125. Oxford: Blackwell.

Perez-Cortes, Silvia, Michael T. Putnam & Liliana Sánchez. 2019. Differential access: Asymmetries in accessing features and building representations in heritage language grammars. *Languages* 4(4). 81.

Peter, Lizette, Tracy Hirata-Edds & Bradley Montgomery-Anderson. 2008. Verb development by children in the Cherokee language immersion program, with

implications for teaching. *International Journal of Applied Linguistics* 18(2). 166–189.

Polinsky, Maria. 2018. *Heritage languages and their speakers*. Cambridge, UK: Cambridge University Press. DOI: 10.1017/9781107252349.

Polinsky, Maria & Gregory Scontras. 2020. Understanding heritage languages. *Bilingualism: Language and Cognition* 23(1). 4–20. DOI: 10.1017/S1366728919000245.

Prévost, Philippe & Lydia White. 2000. Missing surface inflection or impairment in second language acquisition? *Second Language Research* 16(1). 103–133.

Putnam, Michael, Matt Carlson & David Reitter. 2018. Integrated, not isolated: Defining typological proximity in an integrated multilingual architecture. *Frontiers in Psychology* 8. 2212.

Putnam, Michael, Diego Pascual y Cabo & Tanja Kupisch. 2018. Different situations, similar outcomes: Heritage grammars across the lifespan. In David Miller, Fatih Bayram, Jason Rothman & Ludovica Serratrice (eds.), *Bilingual cognition and language: The state of the science across its subfields* (Studies in Bilingualism 54), 251–279. John Benjamins. DOI: 10.1075/sibil.54.12put.

Putnam, Michael T. 2020. One feature-one head: Features as functional heads in language acquisition and attrition. In Pedro Guijarro-Feuntes & Cristina Suárez-Gómez (eds.), *New trends in language acquisition within the generative perspective*, 3–26. Dordrecht: Springer.

Putnam, Michael T. & Birna Arnbjörnsdóttir. 2015. Minimizing (interface) domains: The loss of long-distance binding in North American Icelandic. In Richard B. Page & Michael T. Putnam (eds.), *Moribound Germanic heritage languages in North America: Theoretical perspectives and empirical findings*, 203–223. Leiden: Brill.

Putnam, Michael T., Silvia Perez-Cortes & Liliana Sánchez. 2019. Language attrition and the Feature Reassembly Hypothesis. In Monika S. Schmid & Barbara Köpke (eds.), *The Oxford handbook of language attrition*, 18–24. Oxford: Oxford University Press.

Putnam, Michael T., Lara S. Schwarz & Andrew D. Hoffman. 2021. Morphology. In Silvina Montrul & Maria Polinsky (eds.), *The Cambridge handbook of heritage languages and linguistics*, 613–643. Cambridge: Cambridge University Press.

Riksem, Brita Ramsevik. 2017. Language mixing and diachronic change: American Norwegian Noun Phrases then and now. *Languages* 2(3). 29.

Riksem, Brita Ramsevik. 2018. *Language mixing in American Norwegian noun phrases: An exoskeletal analysis of synchronic and diachronic patterns*. NTNU Norwegian University of Science & Technology. (dissertation).

Riksem, Brita Ramsevik, Maren Grimstad, Terje Lohndal & Tor Anders Åfarli. 2019. Language mixing within verbs and nouns in American Norwegian. *Journal of Comparative Germanic Linguistics* 22(2). 189–209.

Scontras, Gregory, Maria Polinsky & Zuzanna Fuchs. 2015. Heritage language and linguistic theory. *Frontiers in Psychology* 6. Article 1545.

Scontras, Gregory, Maria Polinsky & Zuzanna Fuchs. 2018. In support of representational economy: Agreement in heritage Spanish. *Glossa* 3(1). 1–29. DOI: 10.5334/gjgl.164.

Sherkina-Lieber, Marina. 2015. Tense, aspect, and agreement in heritage Labrador Inuttitut: Do receptive bilinguals understand functional morphology? *Linguistic Approaches to Bilingualism* 5(1). 30–61.

Sherkina-Lieber, Marina. 2020. A classification of receptive bilinguals. Why we need to distinguish them, and what they have in common. *Linguistic Approaches to Bilingualism* 10(3). 412–440.

Siddiqi, Daniel. 2009. *Syntax within the Word: Economy, allomorphy, and argument selection in Distributed Morphology*. Amsterdam: John Benjamins.

Slobin, Dan I. 1973. Cognitive prerequisites for the development of grammar. In Charles A. Ferguson & Dan I. Slobin (eds.), *Studies of child language development*, 175–208. New York: Holt, Rinehart & Winston.

Starke, Michel. 2009. Nanosyntax: A short primer to a new approach to language. *Nordlyd* 36(1). 1–6.

Stump, Gregory. 2017. The nature and dimensions of complexity in morphology. *Annual Review of Linguistics* 3(1). 65–83.

Stump, Gregory T. 2001. *Inflectional morphology: A theory of paradigm structure*. Cambridge: Cambridge University Press.

Svenonius, Peter. 2016. Spans and words. In Daniel Siddiqi & Heidi Harley (eds.), *Morphological metatheory*, 201–222. Amsterdam: John Benjamins.

Tóth, Gergely. 2007. *Linguistic interference and first-language attrition: German and Hungarian in the San Francisco Bay area*. Frankfurt: Peter Lang.

van Hout, Angeliek. 2008. Acquiring perfectivity and telicity in Dutch, Italian and Polish. *Lingua* 118(11). 1740–1765.

Wechsler, Stephen. 2021. The role of the lexicon in the syntax-semantics interface. *Annual Review of Linguistics* 6. 67–87.

Chapter 6

A multi-generational analysis of heritage language complexity

Ashvini Varatharaj[a], Gregory Scontras[b] & Naomi Nagy[c]
[a]University of California, Santa Barbara [b]University of California, Irvine
[c]University of Toronto

Heritage languages are often of interest because of the ways in which they differ from the relevant baseline. Many conceive of these differences as a process of simplification: a loss of inflectional morphology, less lexical richness, etc. Inspired by findings in the literature that decreased complexity in one area of a language may lead to increased complexity in another, we take up the question of whether the changes during the development of heritage languages involve a general simplification, or whether complexity trades off in heritage languages as it does in other languages: as speakers rely less on word-internal structure, word order matters more, and vice versa. We apply information-theoretic measures of complexity in the domain of word structure (i.e., morphology) and word order (i.e., syntax) to six languages from the Heritage Language Documentation Corpus (Nagy 2011), which includes multiple generations of heritage languages and homeland comparators. Our results show partial support for complexity trade-offs in heritage languages, such that as the generations progress, word-structure complexity decreases while word-order complexity increases.

1 Introduction

Languages can be complex in various ways. Existing attempts at objectively quantifying grammatical complexity have primarily focused on morpho-syntactic components, either word-internal structure or the complexity introduced by word order restrictions. Some of these studies calculate complexity on the basis of hand-coded grammatical features (e.g., Shosted 2006, Lupyan & Dale 2010),

Ashvini Varatharaj, Gregory Scontras & Naomi Nagy. 2024. A multi-generational analysis of heritage language complexity. In Maria Polinsky & Michael T. Putnam (eds.), *Formal approaches to complexity in heritage languages*, 133–152. Berlin: Language Science Press. DOI: ??

while others use computational or information-theoretic metrics to calculate the complexity of a grammar based on the behavior that grammar generates (i.e., naturalistic productions from corpora; e.g., Juola 1998, Koplenig et al. 2017). We examine the second type of complexity in this paper.

Importantly, complexity is not a static quantity; the complexity of different grammatical components may shift over time. Complexity also differs across levels of a language. Indeed, a prominent (though problematic) view in language science holds that all languages are equally complex, such that language change involves the redistribution of complexity from one aspect of a language to another. According to this law of conservation of complexity, as it were, complexity can neither be created nor destroyed. While we hesitate to adopt a strong version of this stance (for discussion, see Sampson et al. 2009), we do believe the perspective offers lessons that may help guide inquiry: while total complexity may not be a static quantity, there are likely to be interactions between grammatical components such that increases of complexity in one domain may lead to (or coincide with) decreases in others. Existing investigations have approached these interactions primarily through the lens of idealized monolingualism; the current work investigates complexity in the area of heritage languages.

Heritage speakers are bilinguals who learn their first language (the heritage language) at home. They may then shift to speak the dominant societal language, typically at the onset of schooling (Rothman 2009, Scontras et al. 2015, Polinsky & Scontras 2020). Children usually acquire a heritage language from their parents, who are often recent immigrants. Impressionistically, heritage languages are commonly described in terms of decreased complexity: fewer morphological distinctions, a more limited syntactic repertoire, etc. (see Polinsky 2018 for discussion). Here, we call into question the notion that the process of becoming a heritage language involves only simplification, looking at how complexity changes as a language develops from homeland speakers through successive generations of heritage speakers. Our aim is to understand how (or if) changing complexity in one area of a language (say, word-structure or morphological complexity) interacts with complexity in other areas of the language (say, word-order or syntactic complexity). To accomplish this aim, we use multi-generational data collected as part of the Heritage Language Variation and Change Project (i.e., the largest attempt at documenting multi-generational heritage language productions; Nagy 2009), applying off-the-shelf information-theoretic metrics of complexity to transcribed naturalistic speech from six heritage languages: Cantonese, Faetar, Italian, Korean, Russian and Ukrainian.

In Section 2, we provide background on information-theoretic measures of complexity. In Section 3, we present the complexity metrics we use in more detail,

together with an overview of the data to be analyzed; we then present our results and follow-up analyses in Section 4. We conclude in Section 5 with a discussion of our findings in light of the literature on heritage language and grammatical complexity, noting that our preliminary findings support a trade-off between morphological and syntactic complexity that develops from one generation of heritage speakers to the next.

2 Background

Opinions about language complexity abound. An English speaker trying to track gender on the nouns of Spanish might suspect that the Spanish nominal system features more complexity than English's; a Spanish speaker attempting to internalize the Russian declension system would likely conclude that Spanish's two genders hardly compete with Russian's six cases. But L1 speakers all wind up acquiring the relevant systems, so does any of these languages count as more complex than the others? Despite the abundance of anecdotal intuitions, the trick lies in operationalizing complexity so that it may be subjected to objective – or at least systematic, reproducible – measurement. The first step often involves relativizing complexity to specific aspects of a language or components of its grammar. In what follows, we review several attempts at such an operationalization of complexity.

Our focus is on grammatical complexity, or the complexity of the language-specific knowledge a speaker possesses when they know a language. Other types of complexity focus on the use of the linguistic system; while there are many interesting questions to pose about user-based complexity in heritage languages (for discussion, see Laleko & Scontras 2021), we limit our focus to the complexity of the linguistic system itself. Now, assessments of grammatical complexity are necessarily indirect, owing to the fact that comprehensive descriptions of linguistic knowledge – in other words, fully-specified grammars – have yet to be identified. As is common in (socio)linguistics, we consider performance in communicative tasks as an accessible proxy of this competence. The indirectness enters, then, in the inference of a grammar's content on the basis of the observable linguistic behavior the grammar generates. We may use the observable behavior to construct partial grammars, and then evaluate their complexity via various forms of counting (e.g., Nichols 1992, Bakker 1998, Lupyan & Dale 2010). Or we may skip the partial grammars altogether and directly evaluate the complexity of observed language behavior, with the assumption that the more complex the behavior is, the more complex the grammar that generated it must be. We will

focus on this latter approach, applying complexity metrics directly to naturalistic corpora (i.e., the behavior generated by some grammar(s)) in an effort to assess grammatical complexity.

Corpus-based approaches offer a fast and reproducible method to calculate complexity, with the advantage that there is no need for hand-coding grammatical features of the language under investigation. In other words, corpus-based approaches lessen the need for (and potential bias introduced by) intuitions of the investigator. These methods were initially focused on morphological complexity; they are founded on the idea that morphological complexity depends on the morphological component of a language's generative grammar, which determines the language's inflectional and derivational processes. A productive system will produce different but related word forms (e.g., *walk* and *walked*); as the number of morphological relationships among word forms – and the irregularity of those relationships – increases, so too does the complexity of that system. Information-theoretic metrics can provide us with an understanding of this morphological richness (Gutierrez-Vasques & Mijangos 2020).

Here we review two studies that have inspired our current work. In the first, Juola (1998) pioneers a metric for estimating morphological complexity. Juola identifies complexity with compressibility – or, more precisely, with the lack of compressibility. A more complex string of text will carry more information; it will also be less compressible, because it will require a more complex system to encode and reconstruct (or generate) the original string. This notion of complexity relates to the information-theoretic notion of Kolmogorov complexity (Kolmogorov 1968, Li & Vitányi 2008), where the complexity of a grammar is a function of the length of the minimal description (i.e., specification) of the grammar. Juola recognized that comparison between the amount of morphological information in an original text and in a altered version of the text (without morphological structure) can constitute a measure of the informativeness of the text, and thereby the amount of information encoded morphologically. By estimating the amount of information conveyed by morphology (i.e., word structure), Juola arrives at a method for estimating complexity of the morphological tier of the grammar.

In more detail, Juola artificially inflated the information content of the morphological tier by replacing each word type in a corpus with a unique number. So, *walk* may be replaced by the number 139 and *walked* by 4597. While a compression algorithm can seize on the transparent morphological relationship between *walk* and *walked*, that relationship is destroyed in the case of 139 and 4597; the resulting text with words replaced by numbers is thus less compressible, which

means it is more complex. Inflating morphological complexity becomes more difficult as the morphology becomes less regular; *see* and *saw* have little redundancy for a compression algorithm to seize on, and so replacing the words with random numbers does little to increase complexity.

A ratio of compressed text size between the original text and the morphological-structure-destroyed text gets used as an approximate measure of morphological complexity. We describe Juola's methodology in more detail in the following section; for now, the key to this approach is the fact that, by inflating the information content of the morphological tier, languages with regular, simple morphology will have their information content greatly increased relative to the information in the unaltered text, thus providing a numerical measure to compare morphological complexity across different languages.

Juola applied his metric to Bible translations in six languages: Dutch, English, Finnish, French, Maori, and Russian. Because they communicate the same message, the Bible translations control for overall message complexity; any differences in compressibility across languages are assumed to result from the complexity of message encoding, which an individual language's grammar determines. Juola found that the six languages are ordered as follows with respect to morphological complexity:

(1) Maori < English < Dutch < French < Russian < Finnish

Crucially, the results of Juola's compression-based complexity metric align with those of Nichols (1992), who hand-calculated morphological complexity by counting the number of points at which typical sentences may be inflected in a given language.

In the second study that directly informs our own investigation, Koplenig et al. (2017) implement an entropy-based variant of Juola's metric to estimate both morphological and syntactic complexity, looking for trade-offs between the two quantities across languages in the Parallel Bible Corpus (Mayer & Cysouw 2014). The authors assess intra-lexical and inter-lexical regularities and show that there is a negative correlation between the two: as word-order complexity increases across languages, word-structure complexity decreases, and vice-versa (see also Juola 2008). A particularly interesting result is the observation regarding the diachronic complexity trade-offs within a single language: English. Koplenig et al. use their metrics to show that as English changes from Old English to Middle English, Modern English, and then to English-based Creole, information present in word structure decreases, and word-order information increases. In other words, as the morphology simplifies in the historical development of English such that less information gets communicated via morphology,

the information that would have been communicated via morphology gets carried instead by word order. In our work, we study similar relationships between word-structure and word-order complexity in the development of heritage languages, from speakers in the homeland through subsequent generations of heritage speakers.

3 Methodology

Here we discuss our dataset and the methods we used to assess grammatical complexity.

3.1 Data

The dataset used is obtained from the Heritage Language Variation and Change Project (HLVC; Nagy 2009), which examines usage and change in heritage – that is, non-official – languages spoken in the Greater Toronto Area. The Heritage Language Documentation Corpus (HerLD) documents cross-generational variation in ten heritage languages via digital recordings of spontaneous conversational speech and time-aligned orthographic transcriptions of these conversations. The corpus contains transcriptions of sociolinguistic interviews of about one hour in length and a picture-description task, 10–15 minutes, describing pictures from a children's book called *First 100 Words* (Amery & Cartwright 1987), as well as coded responses to an ethnic orientation questionnaire. For each language, adult heritage language speakers are recorded to represent four generational groups, defined as follows:

Homeland:
- born in the homeland and remained there

1st generation (Gen 1):
- born in the homeland
- moved to the Greater Toronto Area after age 18
- in Toronto at least 20 years

2nd generation (Gen 2):
- born in the Greater Toronto Area (or came from homeland before age 6)
- parents qualify as 1st generation

3rd generation (Gen 3):

- born in the Greater Toronto Area
- parents qualify as 2nd generation

One additional inclusion criterion is that participants self-identify as fluent enough to participate in a one-hour conversation in their heritage language. Data were collected by heritage speakers, recruiting through their personal networks. No fluency or proficiency tests were administered, and participants were selected to represent a diverse array of backgrounds. They vary in terms of education, language attitudes, frequency of language use, etc. More information about the speakers and their selection can be found in Nagy (2011, 2015). Details about data collection are available at https://ngn.artsci.utoronto.ca/HLVC/2_2_linguists.php. For Ukrainian, it has been possible to collect small samples of Gen 4 and Gen 5 speakers. For the Asian languages, there is limited availability of third-generation speakers due to strict immigration restrictions until the 1960's. Table 1 shows the distribution of languages and speakers selected for our analyses.

Table 1: Number of speakers for each language across generations. H: Homeland generation.

	Generation						
Language	H	1	2	3	4	5	Total
Cantonese	-	2	2	2	-	-	6
Faetar	-	9	9	-	-	-	18
Italian	-	5	14	7	-	-	26
Korean	7	3	9	1	-	-	20
Russian	-	11	15	3	-	-	29
Ukrainian	18	3	7	7	3	1	39

3.2 Metrics

Our complexity metric comes from the work of Juola (1998), where the information-theoretic idea of using a compression-based metric for measuring language complexity was proposed. To understand the metric used in this work, one needs to first understand how to quantify information. According to Information

Theory, "information" is the amount of surprisal or unexpectedness present in a text sample – in other words, the extent to which the text is not easily predictable. (If the contents of a text are predictable, then the information it would have conveyed is already in hand, so the text contributes little new information.) The common model to measure information is in the form of transmitting messages (usually treated as stochastic, i.e., randomly determined, or having a random probability distribution or pattern that may be analyzed statistically but may not be predicted precisely) from a source to a listener through some channel. The transmitter encodes the message (in our case, in the form of some text); the listener decodes the message. If the probability distribution for the source producing these messages is known, then we can find an encoding method that transmits the messages with maximum efficiency.[1] This encoding efficiency is known as Shannon's entropy, calculated by the equation $H = -\sum_{i=1}^{n} p_i \log(p_i)$ where p_i is the probability of the ith message being transmitted (given that there are n messages to be transmitted). In real-world scenarios, the exact probabilities are not known but only approximated using different methods (Ziv & Lempel 1977, 1978), which gives an indirect estimation of entropy. Higher-entropy distributions feature greater uncertainty; samples from these distributions are less predictable, which means that encountering one of those samples is more informative.

Next, we must understand what compression means in this context. The process of compression involves removing repeated (or inconsequential) bits of information to make files smaller, typically by reducing the number of bits (1's and 0's) used to store the information (information in this context is represented as 1's and 0's). The main idea of compression within Information Theory is to reduce a communication channel to its maximum efficiency (for example, modern file compression programs use some version of entropy reduction).

Kolmogorov offers another complexity metric that measures a different aspect of informativeness. Kolmogorov complexity measures the information present in a given string of text in terms of the size of the algorithm required to describe that string or regenerate it. The size of the algorithm can be seen as the amount of effort required to transmit the message from source to receiver as assumed in Information Theory. Intuitively, Shannon's entropy can be seen as the upper limit of the Kolmogorov complexity: a decompression program and a compressed file can be used to (re)generate the original message/string. Thus we can say that

[1]By "maximum efficiency" we mean that the system is able to transmit the entire message to the receiver without any loss in information, since the source probabilities are known and lost information can be easily recovered.

a more complex string (in the Kolmogorov sense) will be less compressible and therefore require a larger program and compressed text system to reconstruct. In this way, the compressibility of a text gives us a way to estimate its complexity and the complexity of the grammar that generated it.

For our work, we utilize the gzip compression algorithm. While the algorithm has multiple steps, the most important one for our purposes reduces redundancy. This step involves going through the text and, for each sequence of characters, looking back through the text to see if the same sequence of characters has occurred before. If it has, then the algorithm replaces the current occurrence with a pointer in the current location back to the previous occurrence. In other words, the current occurrence gets deleted and replaced by a pointer, and therefore the file size shrinks.

We calculate Juola's word-structure complexity metric for our heritage data by first replacing each unique word type in the corpus with a random number (using the Keras Tokenizer; Chollet et al. 2015) to destroy any morphological relationship at the word level, thus trying to increase complexity/decrease compressibility. We refer to this version of the text as the "cooked" version of the text (we borrow this terminology from Juola 1998; see Table 2 for an example). As mentioned previously, the words *walk* and *walked* would now be replaced with two unrelated random numbers. While the fact that two words are related and have common characters could be used in the raw text's compression, the relationship no longer exists once we replace the words with numbers, thus destroying the morphological relationship between such related words. For each speaker's text (i.e., the transcription of their descriptions of *First 100 Words* picture book), the original raw text and the cooked text with random numbers are then separately compressed using the gzip compression algorithm. The ratio of these two file sizes is calculated for each speaker, which serves as our metric for word-structure complexity. We refer to this ratio as R/C, where R is the size of the compressed raw text, and C is the size of the compressed "cooked" text.

For an intuitive understanding of the R/C metric of word-structure complexity, consider how the quantity behaves. The compressibility of a cooked file C represents an attempt to inflate the information contained within the morphological tier by destroying relationships across words (e.g., walk :: walked vs. 139 :: 4597). It is easier to artificially inflate morphological information (i.e., increase C) in a language where the morphology does not already contribute substantial information; thus, the ratio of the compressed raw text R to the compressed cooked text C will be smaller (i.e., R/C will decrease). In a language with a rich, informative morphological system, it will not be as easy to inflate morphological information in this way, so C will be smaller relative to R, and thus R/C increases. In practice,

rich morphological systems with large amounts of irregularity – in other words, systems that necessitate a longer description – will yield higher R/C values.

While Juola (1998) presents a way to measure morphological complexity in terms of the information content of a language's word structure, this method can easily be extended to analyze word-order complexity; we simply need a method of artificially inflate the information contributed by word order. Just as replacing word types by integers destroys morphological regularities and thereby inflates word-structure complexity, randomly shuffling the order of words destroys word-order regularities and thus inflates word-order complexity (see Table 2 for an example; cf. Koplenig et al. 2017). To estimate word-order complexity, we randomly shuffle the words in each speaker's data (i.e., raw text) and then compress the resulting file and calculate the ratio between the raw data and the cooked data with shuffled word order, referring to this value as R/C_{shuffled}. The R/C_{shuffled} calculation compares raw to shuffled text; the number-encoded cooked version is not involved.

Table 2: Examples of raw and cooked texts

Raw text	they walked to the station
Number-encoded 'cooked' text (increasing morphological complexity)	13 100 4 210 190
Shuffled 'cooked' text (increasing word-order complexity)	the station they to walked

The intuitive understanding of R/C_{shuffled} mirrors that of R/C: it is easier to artificially inflate the word-order information (i.e., increase C_{shuffled}) in a language where word order does not already contribute substantial information. Thus, in a language with low word-order complexity, C_{shuffled} will be larger and so R/C_{shuffled} will be smaller. In a language with high word-order information, it will be harder to increase the already-high information, and so R/C_{shuffled} will be larger. In practice, R/C_{shuffled} will increase as a language's word order becomes more rigid, since more space will be required to describe the constraints on order (e.g., Ehret & Szmrecsanyi 2019).

We use the correlation between R/C_{shuffled} and R/C, as well as the ratio between the two metrics across generations, to make inferences about the relationship between word-order and word-structure complexity across heritage language generations. To assess the amount of variation in our data, we bootstrap

confidence intervals by re-sampling the data with replacement (Singh & Xie 2008). For each language, we sample (i.e., randomly select) speaker data (with replacement) of size n (where n is the total number of speakers for that language) and calculate the metrics. We repeat this process for a hundred iterations; where 95% of the results fall serves as our (bootstrapped) 95% confidence interval.

4 Analysis and results

Given the sensitivity of information-theoretic measures of complexity to the amount of data under analysis Ehret & Szmrecsanyi (e.g., 2019), we begin by presenting the results for Ukrainian, the language for which we have the most data and thus the best chance of finding reliable results (cf. Table 1). We then expand our analysis to the other languages in our dataset, and in the process we introduce additional analyses to aid in characterizing the pattern of results.

In order to compare more similar content across the heritage languages, we use the speakers' data from the picture-description task for this part of the analysis. Parallel corpora are a valuable resource for many Natural Language Processing tasks and linguistics studies; parallel corpora allow researchers to compare the complexity of message encoding independent from the complexity of the message itself. While not perfectly parallel, the picture-description task data come closer than the sociolinguistic-interview task data to a parallel corpus, given that all speakers were describing the same images.

4.1 Ukrainian

The results of our two complexity metrics are plotted in Figure 1. We see that morphological complexity (R/C) is higher in homeland speakers than in heritage speakers. Among the heritage speaker groups, we further see a trade-off between morphological complexity and word-order complexity: as the heritage generation increases, R/C decreases while R/C_{shuffled} increases; the two quantities are negatively correlated, and the Spearman correlation reaches significance ($\rho = -0.83$, $p < 0.001$). The value of word-order complexity for the homeland speakers is intermediate between the values of generations 2 and 3.

These results suggest that heritage speakers rely less on morphological information to communicate their messages, compared to homeland speakers. And the reliance on morphological information continues to decrease across the generations of heritage speakers. As the information conveyed by word structure decreases, we observe an increase in word-order information among the heritage

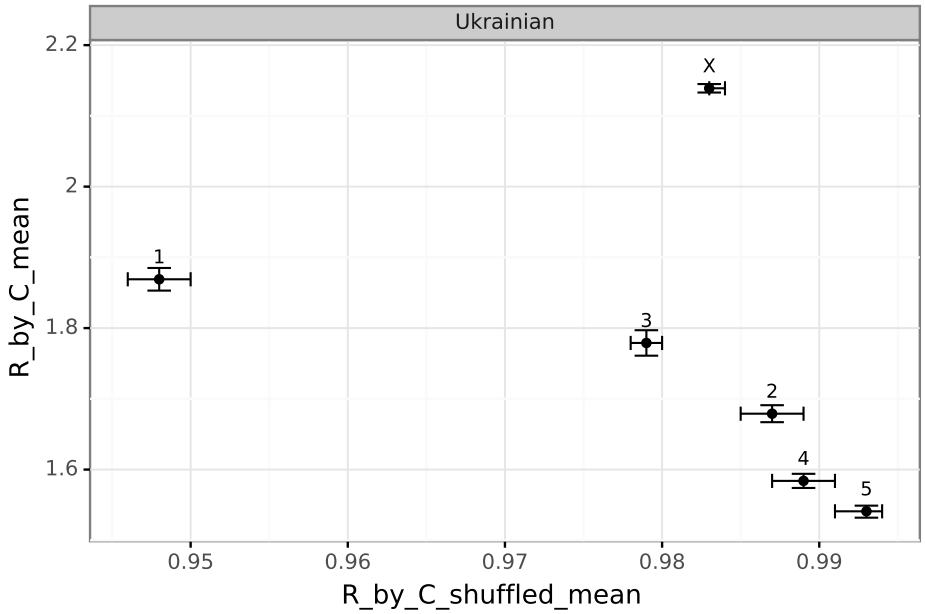


Figure 1: Ukrainian Language complexity metrics: R/C (i.e., word-structure complexity) vs. R/C_{shuffled} (i.e., word-order complexity). Error bars represent bootstrapped 95% confidence intervals

speakers: the information that would have been communicated by word-internal structure presumably gets shifted to information conveyed by word order. This interpretation is confirmed in Figure 2, where we plot the ratio between word-structure and word-order complexity across generations. The complexity ratio should not be confused with the complexity metrics (R/C and R/C_{shuffled}); the complexity ratio is the ratio of the complexity metrics ($R/C \div R/C_{\text{shuffled}}$). As Figure 2 shows, the complexity ratio shifts in favor of word-order complexity as the generations progress away from the homeland ($\rho = -0.94$, $p < 0.001$).

Despite the clear pattern of complexity shift across the generations, the relative order of generations 2 and 3 defies the general trend away from word-structure complexity. While the difference between these two generations is small, it is worth seeking an explanation, as we find traces of this trend in several languages (see Section 4.2). Although the precise source of this pattern remains unclear, a contributing factor could be the change in living conditions in the relevant generations. By the third generation, many families are well-established and able to focus on maintenance of the heritage language alongside English, while the second generation may have felt more pressure to focus resources on English.

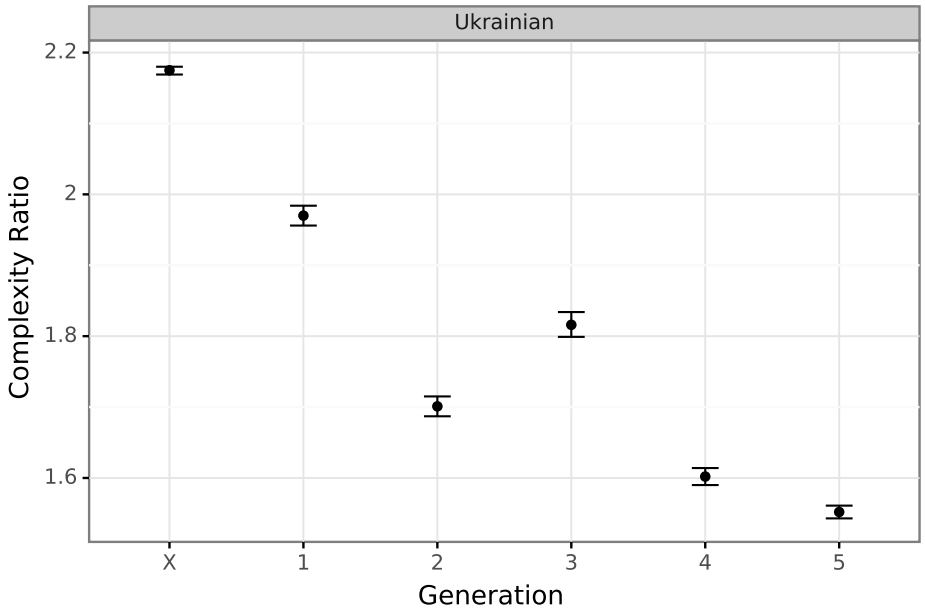


Figure 2: Ukrainian language complexity ratio ($R/C \div R/C_{\text{shuffled}}$) across the generations

Also, in some immigrant communities in Toronto, grandchildren often live with (and thus use their heritage language more with) their grandparents while attending university, suggesting another motivation for similarity between first- and third-generation speech.

4.2 More languages

For the other languages in our analysis, we have fewer data to analyze (cf. Table 1), and so we consider the following results preliminary. Still, some patterns already present themselves.

Figure 3 plots word-structure complexity (R/C) against word-order complexity (R/C_{shuffled}) for the remaining languages: Cantonese, Faetar, Italian, Korean and Russian (as well as Ukrainian). No language clearly replicates the trend found in Ukrainian. Table 3 reports the statistical correlations, repeating the correlation given above for Ukrainian. We see that Ukrainian, Russian, and Faetar have a negative correlation between R/C vs. R/C_{shuffled}, indicating that as word structure increases, information present in word order decreases; these trends are consistent with the findings from Koplenig et al. (2017). However, only in Ukrainian is

the correlation significant – perhaps owing to the smaller datasets for Russian and Faetar. On the other hand, Italian, Korean, and Cantonese have positive correlations, suggesting that as word-structure information increases (or decreases), word-order information also increases (or decreases) – precisely the opposite pattern one would expect on the basis of Koplenig et al.'s findings. However, none of these trends reach significance.

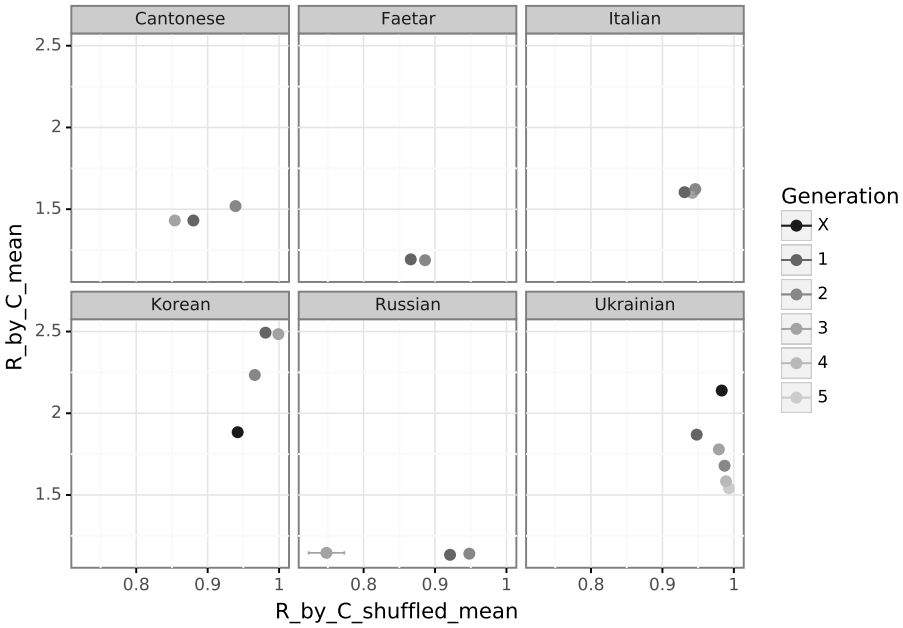


Figure 3: R/C vs. R/C_{shuffled} for all languages

To get a better sense of the changing trend across generations, in Figure 4 we plot the ratio between the two complexity measures across the generations within each language; Table 3 provides the correlations between the generations and the complexity ratio for each language. In addition to Ukrainian, the trend in Italian reaches significance; in both languages we find negative correlations such that, as the generations progress from homeland to later generations, the complexity ratio decreases. In other words, in Ukrainian and Italian, but not the other four languages, we find evidence supporting complexity trade-offs in favor of word-order complexity across generations. While the Italian trend is difficult to read off of Figure 4 given the y-axis scale, the statistics in Table 3 support the reliability of this relationship.

Table 3: Spearman correlation between the complexity metrics (R/C and R/C_{shuffled}), and between generation and complexity ratio. *** indicates $p < 0.001$, ** indicates $p < 0.01$, * indicates $p < 0.05$.

	R/C vs. R/C_{shuffled}	Generation vs. Ratio
Cantonese	0.50	0.50
Faetar	−1.00	−1.00
Italian	0.50	−1.00***
Korean	0.80	0.40
Russian	−0.50	0.50
Ukrainian	−0.83**	−0.94***

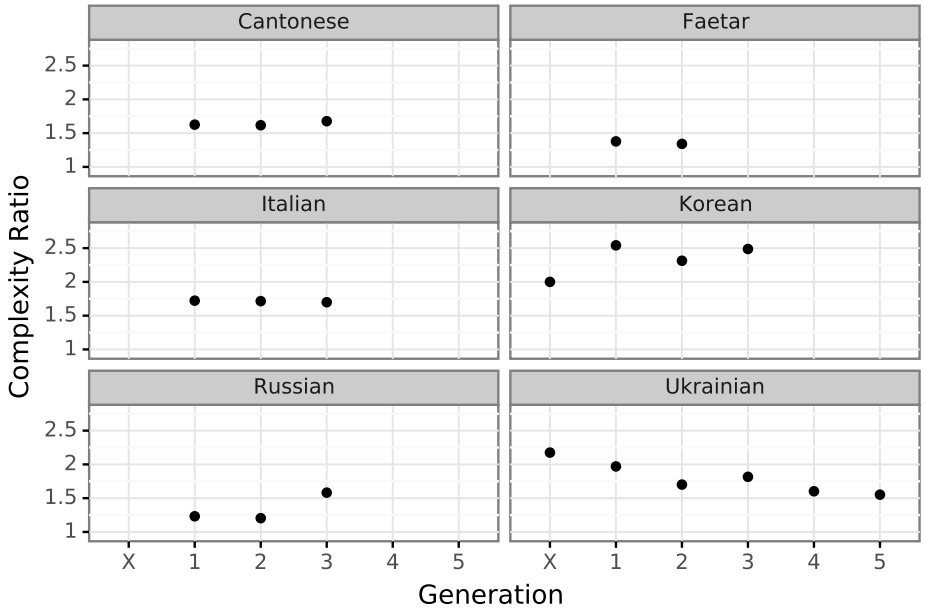


Figure 4: Complexity ratio across generations for all languages: $R/C \div R/C_{\text{shuffled}}$

The ratio analysis also sheds some light on the non-significant positive correlations observed between word-order and word-structure complexity. Although we find trends such that word-order and word-structure complexity increase or decrease with each other (as seen in Figure 3), these increases and decreases do not track the generations of heritage speakers (as seen in Figure 4).

4.3 Followup analysis: Parallel Bible Corpus

Having found clear evidence of a complexity trade-off in Ukrainian and partial evidence in Italian, we decided to explore what might set Ukrainian apart – and in the process also verify the behavior of our metrics. The most obvious explanation for the behavior of Ukrainian vs. the other languages in our study is that we have the most data for Ukrainian, and so it is the only language for which we can get a clear picture of the changing complexity. However, there might be properties of Ukrainian vs. the other languages that incentivize complexity trade-offs in the former. Specifically, it could be the specific language dyad, Ukrainian plus English, that drives the trade-off we observe. As our anecdotal musings in Section 2 illustrate, English is a language with low morphological complexity; Ukrainian is a language with higher morphological complexity. Perhaps the juxtaposition of two systems with drastically different morphological complexity leads to the shifts we observe. To investigate this claim, we applied our metrics to the Parallel Bible Corpus (Mayer & Cysouw 2014) in an attempt to characterize the grammars undergoing change in our heritage speakers.

From the Bible corpora available, we selected English and the languages for which we have heritage speaker data. This process allowed us to analyze Italian, Korean, Russian, and Ukrainian. Faetar and Cantonese were not available in the Bible corpora, but for Cantonese we substituted Mandarin, given the commonalities between the two languages. For each language, we applied our two metrics, R/C and R/C_{shuffled}. Results are plotted in Figure 5. There, we notice that Russian, Ukrainian, and Korean have higher word-structure and word-order complexity than English and that Italian lies close to English. On the other hand, Mandarin (which may be compared to our data for Heritage Cantonese) is less complex than English in terms of word structure and word order. In terms of absolute distance, Ukrainian is farthest from English, and this distance is such that Ukrainian is more complex.[2]

It seems, then, that Ukrainian does stand out from the other languages both in terms of the amount of data we can analyze and in terms of its baseline complexity relative to English. We might wonder, then, whether morphologically-complex languages (relative to the wider community's dominant language in a heritage dyad) result in clearly-observable complexity trade-offs in the heritage varieties. As two systems with very different morphological complexity (e.g., Ukrainian and English) meet in a heritage speaker, the heritage grammar (at least) simplifies its morphology in a way that shifts the communicative burden at least partially to the syntax, such that word-order complexity increases.

[2]These results are consistent with independent analyses of word-order and word-structure complexity (e.g., Bakker 1998, Sadeniemi et al. 2008).

Further investigation would be required to determine if English undergoes an opposite shift for the same speakers.

The wrinkle for this story about the pressures driving complexity change is the comparison between Italian on the one hand and Korean and Russian on the other in our results. In Italian, we observed evidence for complexity trade-offs across the heritage generations; in Korean and Russian we did not. But compared to Korean and Russian, Italian is closer to English in its morphological complexity, so we would expect to observe trade-offs in Korean and Russian as well. However, we had limited data from these languages – including a lack of homeland data for Russian and Italian; as more data become available, it will be important to follow up on this result to see whether the pattern persists.

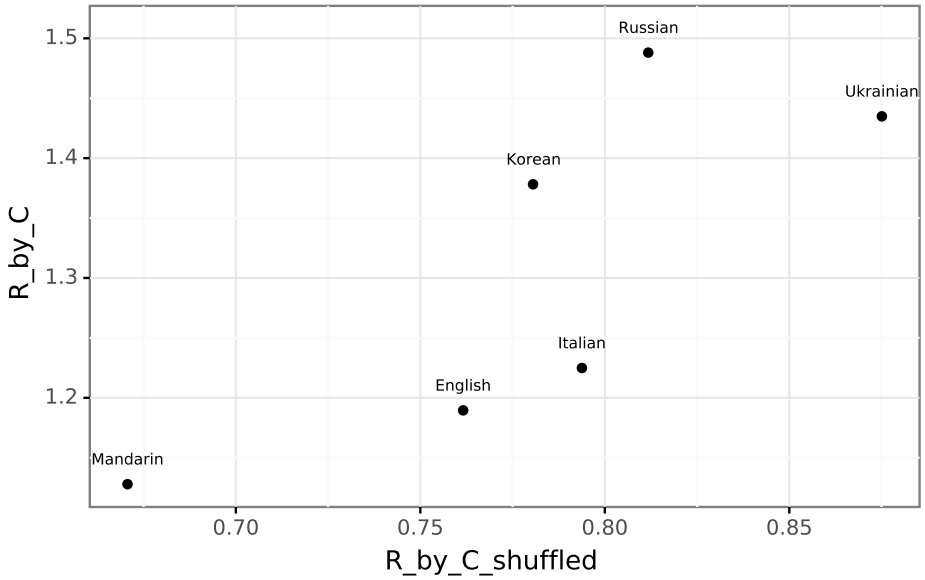


Figure 5: Complexity trade-offs for languages in the Parallel Bible Corpus (R/C vs. R/C_{shuffled})

5 General discussion

Our analysis of word-order and word-structure complexity in heritage languages reveals some support for complexity trade-offs – as operationalized by information-theoretic compression-based metrics – in the development of heritage languages. We saw that in Ukrainian, the language for which we have the most data, complexity trades off across the generations such that, as generational distance

from the homeland increases, word-structure complexity decreases while word-order complexity increases. We found support for a similar trade-off in Italian. The remaining languages in our analysis failed to yield reliable results that track heritage generations.

In an attempt to understand why Ukrainian may stand apart from the other languages in the clarity with which it demonstrates complexity trade-offs, we hypothesized that complexity trade-offs are precipitated by contact between two systems of markedly different morphological complexity. Ukrainian features high morphological complexity, while English's morphology is much simpler; when the two systems come in contact in a heritage speaker who winds up dominant in English, the result is morphological simplification in the Ukrainian grammar. The results of our Bible analysis of the baseline grammars support this interpretation of the results, but we await a clearer picture from Korean and Russian – the two other languages where their morphological complexity relative to English leads us to expect trade-offs.[3] While our dataset's small size is a limitation of our findings, we trust that, as data from multiple generations of heritage speakers become increasingly available, the picture of changing complexity in heritage grammars will become clearer still. Already our results suggest that, rather than being characterized only in terms of a general decrease in complexity relative to the baseline, in heritage languages, as in the languages analyzed by Koplenig et al. (2017), complexity is changing in ways that lead to increases in some areas and decreases in others.[4]

This finding of complexity changes was anticipated by Laleko & Scontras (2021), who discuss the many ways that complexity may exist in heritage languages. We have focused here on morphological (word-structure) and syntactic (word-order) complexity, finding evidence of a trading relationship between the two. Assuming researchers continue to find evidence of such trends across the generations of different heritage languages, it will be important to ask why morphology appears to be so susceptible to change, and why syntax should offer such a ready compensatory mechanism. However, these two notions – morphological and syntactic complexity – do not exhaust the many types of complexity

[3]It would also be interesting to see what happens when the complexity asymmetry shifts in the opposite direction, such that the dominant language features much greater morphological complexity (as in, e.g., a Ukrainian-dominant heritage speaker of English). Unfortunately, such dyads are quite rare in the study of heritage languages (Scontras & Putnam 2020).

[4]We also applied the entropy-based metrics from Koplenig et al. (2017) to our dataset, but we failed to find consistent relationships between word-order and word-structure complexity. We believe our dataset's small sample size could be one reason why we did not observe any clear trends from the entropy-based metrics, which rely on much larger samples to yield reliable results.

that characterize a language and its usage. Grammars are complex systems with interacting components, and there are many other areas where complexity may be shifting (e.g., phonology, pragmatics, or several aspects of usage; see Laleko & Scontras 2021 for discussion). We leave it to future work to explore these other areas, in an effort to arrive at a full picture of complexity in heritage languages.

References

Amery, Heather & Stephen Cartwright. 1987. *First 100 words*. London: Usborne.

Bakker, D. 1998. Flexibility and consistency in word order patterns in the languages of Europe. In A Siewierska (ed.), *Constituent order in the languages of Europe: Empirical approaches to language typology*, 381–419. Berlin: Mouton de Gruyter.

Chollet, Francois et al. 2015. *Keras*. https://github.com/fchollet/keras.

Ehret, Katharina & Benedikt Szmrecsanyi. 2019. Compressing learner language: An information-theoretic measure of complexity in SLA production data. *Second Language Research* 35(1). 23–45.

Gutierrez-Vasques, Ximena & Victor Mijangos. 2020. Productivity and predictability for measuring morphological complexity. *Entropy* 22(1). 48.

Juola, Patrick. 1998. Measuring linguistic complexity: The morphological tier. *Journal of Quantitative Linguistics* 5(3). 206–213.

Juola, Patrick. 2008. Assessing linguistic complexity. In Matti Miestamo, Fred Karlsson & Kaius Sinnemäki (eds.), *Language complexity: Typology, contact, change* (Studies in Language Companion Series 94), 89–108. Amsterdam: John Benjamins.

Kolmogorov, Andrey N. 1968. Three approaches to the quantitative definition of information. *International Journal of Computer Mathematics* 2(1-4). 157–168.

Koplenig, Alexander, Peter Meyer, Sascha Wolfer & Carolin Mueller-Spitzer. 2017. The statistical trade-off between word order and word structure: Large-scale evidence for the principle of least effort. *PLOS One* 12(3). e0173614.

Laleko, Oksana & Gregory Scontras. 2021. On the many dimensions of complexity in hertiage languages. *Heritage Language Journal* 18(2). 1–37. DOI: 10.1163/15507076-12340009.

Li, Ming & Paul Vitányi. 2008. *An introduction to Kolmogorov complexity and its applications*. New York: Springer.

Lupyan, Gary & Rick Dale. 2010. Language structure is partly determined by social structure. *PLOS One* 5(1). e8559. DOI: 10.1371/journal.pone.0008559.

Mayer, Thomas & Michael Cysouw. 2014. Creating a massively parallel Bible corpus. *Oceania* 135(273). Article 40.

Nagy, Naomi. 2009. Heritage language variation and change in Toronto. *Accessed Nov* 3. 2010. http://ngn.artsci.utoronto.ca/HLVC.

Nagy, Naomi. 2011. A multilingual corpus to explore geographic variation. *Rassegna Italiana di Linguistica Applicata* 41(1-2). 65–84.

Nagy, Naomi. 2015. A sociolinguistic view of null subjects and VOT in Toronto heritage languages. *Lingua* 164. 309–327.

Nichols, Johanna. 1992. *Linguistic diversity in space and time*. Chicago/London: University of Chicago Press.

Polinsky, Maria. 2018. *Heritage languages and their speakers*. Cambridge, UK: Cambridge University Press. DOI: 10.1017/9781107252349.

Polinsky, Maria & Gregory Scontras. 2020. Understanding heritage languages. *Bilingualism: Language and Cognition* 23(1). 4–20. DOI: 10.1017/S1366728919000245.

Rothman, Jason. 2009. Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International Journal of Bilingualism* 13(2). 155–163.

Sadeniemi, Markus, Kimmo Kettunen, Tiina Lindh-Knuutila & Timo Honkela. 2008. Complexity of European Union languages: A comparative approach. *Journal of Quantitative Linguistics* 15(2). 185–211.

Sampson, Geoffrey, David Gil & Peter Trudgill. 2009. *Language complexity as an evolving variable*. Oxford: Oxford University Press.

Scontras, Gregory, Zuzanna Fuchs & Maria Polinsky. 2015. Heritage language and linguistic theory. *Frontiers in Psychology* 6. 1545. DOI: 10.3389/fpsyg.2015.01545.

Scontras, Gregory & Michael T. Putnam. 2020. Lesser-studied heritage languages: An appeal to the dyad. *Heritage Language Journal* 17(2). 152–155.

Shosted, Ryan K. 2006. Correlating complexity: A typological approach. *Linguistic Typology* 10(1). 1–40.

Singh, Kesar & Minge Xie. 2008. *Bootstrap: A statistical method*. Unpublished manuscript, Rutgers University, USA. http://www.stat.rutgers.edu/home/mxie/RCPapers/bootstrap.pdf.

Ziv, Jacob & Abraham Lempel. 1977. A universal algorithm for sequential data compression. *IEEE Transactions on Information Theory* 23(3). 337–343.

Ziv, Jacob & Abraham Lempel. 1978. Compression of individual sequences via variable-rate coding. *IEEE Transactions on Information Theory* 24(5). 530–536.

Chapter 7

Non-monotonic functional sequences: A new metric for complexity in heritage languages

Roberta D'Alessandro[a] & Silvia Terenghi[a]
[a]Utrecht University

This paper presents some evidence that language change in heritage languages (and beyond) systematically responds to general factors of language design when it comes to fixed sequences of functional heads within given domains. Concretely, we investigate patterns of change across various heritage languages, both in the word-internal domain (person and number features) and at the sentence level (word order): we show that change in these different domains is consistently shaped by a bias towards monotonicity and uniformity in computation, such that points of non-uniformity in the relevant sequence can be predicted to be the gateway to change. Crucially, this change systematically brings about a reduction in complexity; as such, these factors are proposed as a new metric for linguistic complexity.

1 Introduction

Complexity is a recurrent concept in the analysis of heritage grammars. Nonetheless, a rigorous, agreed upon definition of complexity is currently lacking. In this paper, we restrict the focus to a specific set of purely syntactic phenomena, the derivation of which can be taken to hinge on feature sequences. We show that, when fixed sequences of functional heads are concerned, complexity can be understood as a correlate of properties inherent to this sequence, and more specifically to their values. We identify two general factors of language design: (i) bias towards monotonicity and (ii) uniformity in computation; building on these, we show that it is possible to predict that points of inconsistency across the relevant

Roberta D'Alessandro & Silvia Terenghi. 2024. Non-monotonic functional sequences: A new metric for complexity in heritage languages. In Maria Polinsky & Michael T. Putnam (eds.), *Formal approaches to complexity in heritage languages*, 153–180. Berlin: Language Science Press. DOI: ??

feature values (from + to − and from − to +) will be the gateway to change in heritage languages, as well as in other forms of change, most notably spontaneous, diachronic change.

We first provide an overview on complexity in general and in heritage languages (HLs henceforth) in particular (Section 2). We further argue that the concept of complexity needs to be related to the monotonicity profile of the relevant functional sequence. With this background in place, we explore the proposal by considering two domains, each related to one specific sequence of elements. Section 3 discusses phenomena at the word-internal level, where the hypothesis is illustrated by means of heritage grammars that display change in the person and number domains. In Section 4, instead, the hypothesis is illustrated by considering the phenomena at the sentence level, and more specifically word order facts as found across heritage languages. Section 5 concludes.

2 Complexity in heritage languages

Heritage languages (HLs) are defined in different ways depending, among other things, on various linguistic traditions. Minority languages that are in balanced or displacive contact with a majority language spoken in a given area, as well as dialects or variants of the same language, and languages spoken by emigrants' children, all fall into the category of HLs.[1] In this paper, we will use the tag HL to refer mostly to those languages spoken by the children of immigrants, learned in a naturalistic environment, for instance at home or within a small community, but crucially different from the dominant/official language(s) spoken in the larger community of which they are part. Quoting Polinsky (2018: 9), "[a] heritage language speaker (for short, heritage speaker) is a simultaneous or sequential (successive) bilingual whose weaker language corresponds to the minority language of their society and whose stronger language is the dominant language of that society".

The study of HLs has developed in different directions in the last few years: on the one hand, the focus has been put on the divergence of HL grammars from so-called baseline grammars (see at least discussion in Polinsky 2018: 1.3.3 for the concept of divergent attainment, Pires & Rothman 2009 for that of missing-input

[1]According to the typological tradition of contact studies, *balanced contact* obtains "in a situation of a long-standing linguistic area and stable multilingualism without any dominance relationships" (Aikhenvald 2006: 42); *displacive contact* happens instead "if one group aggressively imposes its language on another group, [resulting] in language displacement, loss of the language's own features, and, ultimately, language shift" (Aikhenvald 2006: 43).

competence divergence, and Montrul 2008 for that of incomplete acquisition); on the other hand, the focus has been put on the speaker's mastery of the language processes (starting from the Shallow Structure Hypothesis by Clahsen & Felser 2006, and especially studies involving interfaces Sorace 2011). Yet another path is taken by studies, like Bayram et al. (2019), who take into account the role of inter-generational attrition in HL acquisition, and put emphasis on the fact that "divergent" attainment could be due to qualitatively different input with respect to that to which monolingual learners are exposed.

This paper takes a slightly different viewpoint, by focusing exclusively on the grammatical system of HLs, putting aside all considerations on performance, onset, fluency, number of languages spoken and their order of acquisition. Our aim is to identify general principles that govern HLs and constrain the ways in which they deviate from the original language. In other words, our aim is to discuss some principles of language change, where language is intended as grammar.

2.1 The problem of the baseline

Whether the focus is on the grammar or on the speaker, HLs have typically been tackled in a comparative fashion: how has a given HL changed (i.e., how has it become simpler or more complex) with respect to its baseline? And what is this change due to?

When trying to understand the mechanisms behind language change, the first problem is to define the system against which grammatical change can be assessed. This is a well-known issue, usually referred to as the "baseline problem" (for an overview of the baseline problem, see Polinsky 2018: 1.1.2, Aalberse et al. 2019: Ch. 6, Bayram et al. 2019, D'Alessandro et al. 2021). Identifying the baseline is not an easy task, especially when dealing with minority and/or non-standardized languages. For example, is the baseline for Heritage Italian spoken in the US the Italian spoken in Italy today? Or is it rather the language to which the heritage speakers were exposed during acquisition? The "deviating value" which appears to be the result of language change might have already been present in the baseline, for instance because the original variety was not standardized and presented wide microvariation. If this original microvariation is not documented to start with, identifying change in non-standardized varieties becomes nearly impossible. A further issue regards the fact that HLs are often compared to their counterparts in the language homeland in their contemporary versions, and not, for instance, to the varieties that were spoken in the country of origin at the moment in which the emigrants left them.

The problems just mentioned make the issue of identifying change more difficult to tackle in the absence of clear understanding of the underlying mechanisms of change. While HL studies are a subset of bilingualism studies, we need to understand the underlying mechanisms guiding change and constraining it. In the absence of a clear view on such mechanisms, we are left with no guidance as to what is possible (and can possibly have been the result of language change) and what not. A similar *desideratum* has been recently expressed by Polinsky (2022): describing some phenomena that underwent change, "itemizing tokens of change" so to say, is not going to bring us too far, if we do not identify some general laws governing said change. These laws can help us solve the issue regarding the possible input for a given phenomenon, in the absence of empirical evidence indicating where the change started from.

In this paper, we will present one such underlying law, identified not only on the basis of HLs, but also on the basis of diachronic evidence. This law, which we call the *monotonicity bias*, seems to inform language change in contact as well as in diachrony. We will present some case studies, both at word and at sentence level, focusing on HLs. More concretely, we will argue that HLs tend to simplification, but only in those areas of language directly related to grammatical features. Before moving on to the discussion of the case studies, we briefly touch upon the definition of simplification and complexification in language change.

In fact, one of the tendencies that have been pointed out for HL grammars is that they are significantly less complex than their corresponding baseline (see Polinsky & Scontras 2020 for discussion).[2] This has been attributed to different factors, including: HL speakers tend to avoid ambiguity/indeterminacy (Polinsky 2018: 5.2; the "ambiguity problem" in Polinsky & Scontras 2020), for instance by avoiding polyfunctional words; they avoid silence (the so-called "silence problem": Laleko & Polinsky 2017; Polinsky 2018: 6.5). However, from a strictly grammatical viewpoint, it is not obvious what this "simplification" amounts to, or whether we can talk of simplification at all, in the first place. Here, we will not use the term "simplification" with respect to performance-based phenomena: as stated above, we will focus exclusively on structure. To do this, we need to briefly discuss the concepts of complexity and markedness; subsequently, we will outline a system predicting functional feature-related change and present evidence for it.

[2]From now on, we will simply refer to the baseline as the system against which we observe change, bearing in mind what was discussed in the beginning of the section.

2.2 Complexity and markedness

Complexity is an elusive concept. If the aim is to determine whether complexity has increased or reduced in a system, one needs to have a way to quantify complexity, to measure it. There have been attempts to measure complexity, both for HLs (for an overview, see Varatharaj et al. 2024 [this volume]) and for languages in general.

One thing to bear in mind is that, although complexity and markedness are obviously very different concepts, complexity has often been reduced to markedness, with the underlying assumption that more marked elements are more complex. Observe that, while markedness traditionally refers to one item or one paradigm, and is determined as a difference with respect to the rest of the system, complexity usually refers to the system as a whole, and is determined by means of comparisons between systems. Markedness of several forms in a system can result in the system being more complex, for instance. This idea has been informed by the same observation regarding the decrease of complexity in diachronic linguistics: languages tend to eliminate complexity through time, and marked elements are also eliminated by the system through time. Something similar has been claimed, on different channels, in contact studies, for instance those on creolization (see for instance McWhorter 2001, among many others).

The correspondence between complexity and markedness is somehow intuitively right, but it suffers from some flaws that have been highlighted by many, among which Haspelmath (2006). In a qualitative fashion, Haspelmath underlines that markedness has different meanings when related to different aspects, and that simplification in one area can mean complexification in another (in this respect, his conclusions are not different from those in Varatharaj et al., this volume). Consider for instance a clitic-left dislocation construction in Standard Italian, like the one in (1a):

(1) a. La torta l'hai mangiat-a
the cake it=have.2SG eaten-PTCP.F.SG
'The cake, you ate it.'

b. Hai mangiat-o la torta
have.2SG eaten-M.SG the cake
'You ate the cake.'

(1a) is quite transparent from a discourse viewpoint, with the object appearing first in the sentence, which makes it immediately clear that one is talking about a cake, the topic. However, if we look at syntactic complexity in terms of number of syntactic operations, the situation is reversed: the object is left-dislocated

(which means that a movement operation is required); this object dislocation also triggers agreement, which is absent when the object is *in situ*. Simplification in interpretation and understanding of discourse corresponds to complexification in syntactic operations, very often also reflected in slowness of processing because of the establishment of a dependency between the object and the clitic (see for instance Sequeros-Valle et al. 2020). This means that we need to identify not just a measure of complexity, but also the domain in which complexity is assessed.

Regarding the idea of exploiting markedness to identify complexity, it needs to be recalled that the concept of *markedness* was first introduced by Trubetzkoy and Jakobson in the 1930s (Trubetzkoy 1931, 1939, Jakobson 1932, 1939), and was mainly used to refer to the characteristics of a grammatical item. As an example, consider the voiced/voiceless alternations in consonants: voiced consonants have, according to Jakobson, an additional "specification" compared to voiceless consonants. From this perspective, "while the optimal consonant is voiceless and the optimal vowel is voiced, the voicing of consonants and, in very rare instances, the unvoicing of vowels, may be utilized as one of the various phonetic attenuations of the maximum contrast CV" (Jakobson & Halle 1956: 56–57). According to this line of thought, that the voiceless consonants are unmarked is also shown by the existence of final devoicing rules, "erasing" the voice/marked feature, in languages like Russian. The markedness on one item has been exploited very often to investigate "morphological complexity".

An example of morphological markedness which is difficult to master for HL speakers can be the Italian finite verb inflectional morphemes, which encode information about the person and number of the subject, the tense, aspect and mood of the verb. These morphemes are semantically marked, since they contain many meanings, and also morphologically marked, because these meanings that are mapped to one exponent simultaneously are not immediately identifiable as the morphology is often "irregular". HLs tend to move in the direction of simplification of semantic complexity in the inflectional system. In a recent study, Andriani & D'Alessandro (2022) show that the inflectional system of Italo-Romance HLs in the Americas is heavily reduced: speakers pick one of the two strategies: they either replace the inflected form with a default one (like the 3rd person singular form of the present tense) or they delete the auxiliary altogether. A similar process is found in creole languages like Papiamento, where the auxiliary only encodes tense (*ta* for the present tense; vs *a* for the past tense) but no *phi*-agreement features.

Morphological markedness can also be tackled from a paradigmatic viewpoint, for instance by isolating the verbal paradigm in a language L and checking how many overt inflectional forms it includes. The more inflectional forms a paradigm

contains, the more marked it will be. The more inflecting categories a language has, the more complex (at least, morphologically) the language is. This position is adopted by Nichols (1992), who defines complexity in terms of inflecting categories in a language. Observe that if the mapping between the exponent and its meaning is a bijective function, this does not indicate necessarily more complexity: the system is richer, but transparent, therefore not necessarily harder to master.

In the remainder of the chapter we will build on intuitions related to morphological complexity, but we will not be adopting this approach to quantify complexity. Rather, we propose a definition of complexity based on morphological structure, taking features and the sequence of functional heads which constitute words as its primitives, in conformity with our task, i.e. to identify an underlying principle governing complexity and simplification in language change. More concretely, we put forth one such underlying principle, namely what we call the *monotonicity bias* (Terenghi 2021a, 2022b, 2023: 173 ff.). We will base our analysis on the kind of markedness which Haspelmath (2006) dubs as "markedness as default from parametric settings", stemming from Chomsky & Halle (1968: Ch. 9) and Kean (1975). The basic idea is that markedness is given by "the odd one out" with respect to a system. These diagnostics could be easily put to use to identify the mechanisms of language change, in at least two ways. The first way would be to actually count the number of irregular words in a language, and check whether they are systematically replaced by regular forms. This seems to be the case in HL, according to what is reported by Aalberse et al. (2019). The second way would be to extend these considerations to all grammar modules, and for instance establish a correspondence between portmanteau morphs at morphological level and complex functional heads, encoding more than one piece of semantics, at syntactic level. Consider again the tense head in Romance: this head is considered to encode at the same time tense, aspect, mood, and *phi*-features. It is a "portmanteau" functional head, which parallels its morphological counterpart. Does some of the information on this head tend to disappear, or does it become more inconsistently marked, or does it settle on a reduced form? Does subject agreement disappear, or does it reduce? Does mood disappear? All these questions have been posed in HL studies, and have been given positive answers: see, for instance, van Osch & Sleeman (2018) on the disappearance of subjunctive in heritage Spanish spoken in the Netherlands. While we do seem to have collected quite a large amount of evidence in favour of simplification of functional heads, the principle underlying this simplification is still obscure.

To understand what this means, we borrow an example by Roberts & Holmberg (2010) on word order in Japanese, a head-final language. Under the assump-

tion that head-finality is reached through movement to the specifier of dedicated functional heads, and that movement is a costly operation (contra Chomsky 2013 etc.), as well as from a processing/interpretational viewpoint), a verb-final language should be more marked than a verb-initial language. The standard assumption regarding movement in minimalist syntax is that it is triggered by an EPP-feature on functional heads, which attracts the moving XP to the specifier of the head featuring the [EPP]. Considering this, head-final languages should be very marked, as they would need an "extra" EPP-feature on every head. Languages like German or Latin, with mixed word order, might be considered as partially marked, given that not all heads would require the EPP-feature, and harmonically head-initial languages like Italian or English would be unmarked.

Building on Chomsky & Halle (1968), however, Roberts & Holmberg (2010) reconsider markedness not as arising from the presence of an additional feature on one functional head but, more holistically, as resulting from a deviation from the paradigm, which in the case of syntax is identified in the set of all functional heads of a language. Through this lens, harmonic head-final languages like Japanese are not marked systems at all, as every functional head carries an EPP-feature. Likewise, consistently head-initial languages are not marked, as no functional head carries an EPP-feature. The only marked systems are those that give origin to disharmonic word orders, like that of German for instance, where head-finality and head-initiality are both present in the system because the EPP-feature is present only on some functional heads, and its distribution depends on whether the clause is a root or embedded one.

Roberts & Holmberg (2010) conclude that the preference for harmonic ordering seems to derive from an overriding tendency for independent parameters to conspire to produce a certain type of grammar. This intuition had already been expressed, in functional terms, by Hawkins (1983) under the notion of cross-categorial harmony, whereby languages are preferred if their constituents display a harmonic ordering. Roberts & Holmberg (2010: 39–40) rephrases this proposal as in (2a) and, more formally, in (2b):

(2) a. There is a preference for the EPP-feature of a functional head F to generalise to other functional heads G, H ...

b. For a class of heads H, *u*EPP for $H_{F:\_} \neq v \rightarrow \left\{ \begin{array}{l} [+\text{EPP}]/v_{+\text{EPP}}; \\ [-\text{EPP}] \text{ elsewhere} \end{array} \right\}$

In summary, for some languages, the unmarked value for the functional heads is [+EPP]. These are OV-languages, where the object is attracted across the verb to the specifier of a higher functional head. For some other languages, the unmarked value for the probe heads is [−EPP]. These are the VO-languages. Mixed systems,

disharmonic word orders, where the EPP-feature is present on only some heads, are more marked (see also Biberauer & Sheehan 2013).

Roberts & Holmberg (2010: 41) propose that this tendency toward uniformity is also the driving principle behind the acquisition of word order. Learners exploit pieces of input, focusing their attention in particular on the *v* head, which is the core head in a clause as it encodes transitivity. Once the [EPP] value on *v* is identified, it gets transferred to all the other functional heads in grammar, according to a mechanism called the Generalization of the input:

(3) Generalization of the input (Roberts & Holmberg 2010: 41)
If acquirers assign a marked value to H, they will assign the same value to all comparable heads.

Our analysis of complexity will build heavily on (Roberts & Holmberg 2010)'s intuitions: we will show that language change in contact, in particular in HLs, tends to unmarked systems, i.e. systems that only include functional heads with the same value. We will call them *monotonic*. Before delving into monotonicity, a last methodological note is required. An observation borrowed from classical phonological studies, like Chomsky & Halle (1968) or Kean (1975), is that markedness arises "one step at a time". In phonological terms, for instance, starting from the fact that all languages have an /a/, the next step will be to add height, and therefore there will be systems with /a/, /i/, and /u/; then, anteriority will be added, and so on, but crucially markedness will not jump ahead and skip one of these steps. We observe the same mechanism in the uniform restructuring of functional sequences: it will be shown that this kind of simplification takes place one step at a time, along the functional sequence.

2.3 (Non-)monotonic functional sequences

We first start by assuming that whenever the derivation of a given phenomenon can be shown to exclusively hinge on an underlying sequence of features, properties inherent to that sequence will determine the complexity level for the given phenomenon: more specifically, sequences of features that only include functional heads that share one and the same value (harmonic, in Roberts & Holmberg' terms) are regarded as less complex than sequences of features that include functional heads with different values (disharmonic). Distinct from Roberts & Holmberg, however, we extend this proposal to the word-internal level, too. This allows us to find a previously unnoticed parallelism across the nominal and the clausal domain, as in both cases the gateway for change within the system can be shown to be exactly the point of the sequence where the relevant features switch

values. We refer to sequences that involve such switches as non-monotonic sequences and propose that they are more complex due to a general bias towards monotonic computation.

Monotonicity is a property whereby (mathematical) functions do not vary in tone for a given values interval. Monotonic functions are either entirely non-decreasing (monotonically increasing functions) or entirely non-increasing (monotonically decreasing functions); functions that instead are partly increasing and partly decreasing can be defined as non-monotonic. The general rationale behind the notion is that, given a partial order within a domain, monotonicity consistently preserves it or reverses it. In this sense, monotonicity has been shown to be also relevant beyond the mathematical level, for other cognitive modules. Interestingly, monotonicity shapes language, too, and it most famously does so in the domain of quantifiers. This line of research goes back at least to Barwise & Cooper 1981, where the relevant partial ordering is the one that exists between two sets that stand in a subset–superset relation.

In this paper we extend this notion to the *phi*-features domain in syntax (based on Terenghi 2021a, 2022b, 2023). We assume an action-on-lattice semantics for person and number features (Harbour 2008, 2014, 2016; see discussion in Section 3.1), whereby features denote sets and their values (+ and −) denote operations performed on these sets. Hence, under the assumption that person and number features denote sets and that these sets are further nested (for instance, the author is a subset of the participants which are a subset of person: author $\subseteq$ participant $\subseteq$ π), the notion of partial orderings becomes relevant in this domain, too. Crucially, this ordering is consistently preserved or reversed by sequences of one and the same operation (i.e. sequences of + or of −), but is obliterated if the sequence of features carry mismatching values (+ and −).

Bridging the gap with the discussion in Section 2.2, Biberauer (2017, 2019) proposed monotonicity (intended as uniformity, and not in its technical meaning) as a general principle of language design, whereby languages (and more specifically: language learners) "generalize over as large a domain as possible to create formally defined domains sharing a particular property" (Biberauer 2019: 69); this proposal is supported by word order facts, and specifically by the derivation of the Final-over-Final Condition (for which, in general, see Sheehan et al. 2017) and by the notion of Phrasal Coherence that constrains nominalisations and verbalisations.

In this paper, we provide new evidence for a bias towards monotonic computations by illustrating how it applies to change in heritage languages, both at word-internal and at sentence level; note, however, that by virtue of the cognitive underpinnings of this metric for complexity, the monotonicity bias is in

fact expected to apply to language (and hence language change) more widely. This is supported by parallelisms between change in heritage varieties and in the diachronic endogenous development of those same varieties; however, this discussion exceeds the scope of the present chapter (for a comparison between language change in heritage and in diachrony, see D'Alessandro et al. forthcoming(b)). Instead, the next two sections put to the test these hypotheses on complexity in the heritage domain. Section 3 introduces examples of word-internal feature sequences, where higher complexity ultimately triggers feature loss, accounting for some change patterns attested in heritage languages. Section 4 turns instead to examples of parameter hierarchies, with ramifications relative to sentence-level facts, and focuses in particular on word-order issues in heritage languages.

3 Sequences at word-internal level

In this section, we consider change as attested in heritage person and number systems. Our preliminary assumption is that the person and number domains can be construed as being yielded by a sequence of features merged in the functional spine of the relevant elements (personal pronouns, demonstrative forms, nouns, *etc.*), as swiftly reviewed in Section 3.1. Granting this, we explore the patterns of change attested in demonstrative systems in heritage Italo-Romance varieties (first-hand data; Section 3.2) and in number systems in heritage Semitic varieties (data from the literature; Section 3.3). With these case studies, we show that if the relevant feature sequence is non-monotonic, the overall system is unstable and the category that is non-monotonically derived is progressively lost.

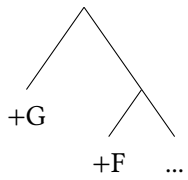
3.1 Sequences of action-on-lattice features and monotonicity

The functional sequence that we consider at word-internal level is the one yielded by the sequence of features that are active in the derivation of a given form; for instance, given a personal pronoun, we are concerned with the set of person features involved in the derivation of that pronoun.

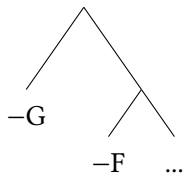
For the present discussion we only consider person and number features. Importantly, we regard these features as denoting sets; by means of their values (plus +, or minus −), these sets are the basis for operations (addition and subtraction) on a further set, which syntactically is their complement (this latter set might itself be the result of earlier feature operations), in line with the action-on-lattice features framework set by Harbour (2008, 2011, 2016), among others.

By performing such operations, features partition the input set (that denoted by their complement) in different subsets: each of these subsets is identifiable with a person or a number category, depending on the features involved. It follows that the ordering of operations is important for the derivation (set-theoretic operations may be non-commutative; see Harbour 2016: 66 for discussion). For the purposes of this article, we implement this under 1 Feature–1 Head assumptions: concretely, we regard each feature as a head; thus, the set of features is scattered along the functional spine and the ordering of operations can be straightforwardly read off the tree (see Terenghi 2023: 92–93 for discussion; but see Harbour 2016 for a different implementation). As such, the functional sequence under investigation at the word-internal level is modelled as a sequence of positively and/or negatively valued features, as illustrated in (4) in an abstract fashion (F and G are features):

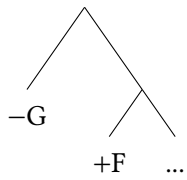
(4) a.



b.



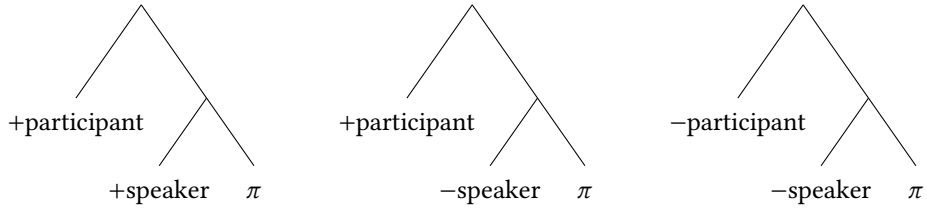
c.



Against this background, we propose that derivations such as the one instantiated by (4c) are more complex, by virtue of their non-uniform sequence of feature values. The rationale behind this hypothesis is that, if feature values are indeed taken to denote different of (set-theoretic) operations, sequences of consistently positive features ([+G] ≻ [+F]; (4a)) and sequences of consistently negative features ([−G] ≻ [−F]; (4b)) constitute sequences in which one and the same operation is reiterated; instead, sequences that include both positive and negative features ([+G] ≻ [−F] or [−G] ≻ [+F]; (4c)) must involve two different operations. These latter sequences can be flagged as being more complex due to a third-factor rooted monotonicity bias (Terenghi 2021a, 2023), in line with the discussion presented in Section 2: that is, grammar favours monotonic sequences (where one and the same operation is reiterated, as denoted by harmonic sequences of feature values) and disfavours non-monotonic sequences (where two different operations are reiterated, as denoted by non-harmonic sequences of feature values).

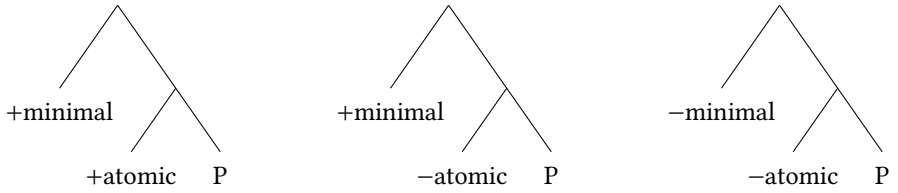
In what follows, we investigate how heritage speakers treat non-monotonic functional sequences in the person and number domain; therefore, we will only be concerned with baseline systems that make three-way oppositions (derived

by the activation of two features: at least one category needs to be derived by a non-monotonic sequence of features) and will leave aside smaller systems (a two-way opposition can be derived by the activation of a single feature: [±F]). More concretely, we will explore ternary demonstrative systems for the person domain (i.e. systems that include a "that/there near you" form) and ternary number systems for the number domain (i.e. systems that include a dual form). The relevant featural derivations assumed in what follows (based on Harbour 2016 and Harbour 2014, respectively) are shown in Figures 1 and 2.



(a) 1st person (b) 2nd person (c) 3rd person

Figure 1: Ternary person systems



(a) Singular (b) Dual (c) Plural

Figure 2: Ternary number systems

In the interest of space, the structures in Figures 1–2 will be simply reproduced as sequences of functional applications; for instance, Figure 1a will be represented as, where each set of brackets represent successive functional applications:

(5) +participant(+author(π))

The monotonicity bias predicts that the featural sequences in Figures 1b and 2b, i.e. those that involve both feature values (+/−), will be more complex and as such prone to change. As the next two sections show, this prediction is borne out.

3.2 Demonstrative systems in heritage Italo-Romance

In this first case study, we consider exophoric demonstrative data from two heritage southern Italo-Romance varieties: heritage Sicilian and heritage Abruzzese. Exophoric demonstratives denote the location of a given referent with respect to a deictic centre (Lyons 1977, Diessel 1999, a.o.), which can be identified with at least one of the discourse participants. Most typically, the deictic centre coincides with the speaker, as is the case for English (if a referent is located about the speaker, it will be denoted by *this*; if it is far from the speaker, by *that*); but there is cross-linguistic variation in this respect, and systems are also attested that encode either proximity of a given referent to the either or both discourse participants (*this near the speaker and/or hearer* vs. *that far from the participants*; see e.g. Catalan or Brazilian Portuguese), or that contrastively encode proximity of a given referent to the hearer alone (*this near the speaker* vs. *that near the hearer* vs. *that far (from both)*). The homeland counterparts (used here as baseline; see remarks in Section 2.1) of the two heritage varieties under discussion display a system of the latter type, as illustrated by Eastern Abruzzese in (6):

(6) Eastern Abruzzese demonstrative system (ternary)
   a. queʃtə
      'this near me'
   b. quessə
      'that near you'
   c. quellə
      'that far from us'

As observed by Terenghi (2022a), speakers of heritage Abruzzese and heritage Sicilian tend to lose the contrastive encoding of one of the three original domains, and more specifically the hearer-related one, regardless of the deictic structure of the demonstrative systems in the dominant varieties. This was assessed by means of both comprehension and production tasks (picture-sentence matching task and guided production, respectively); the results, taken from Terenghi 2022a: 9 are reproduced in Figure 3.

Crucially, as Figure 3 shows, the semantic domain that invariably undergoes loss is the hearer-related one, that is: the only one derived by a non-monotonic feature sequence. In fact, Figure 3 highlights a stark contrast between the latter and demonstrative forms reducible to 1st and 3rd person (i.e. the monotonically-derived person categories), which are interpreted and produced in a target-like

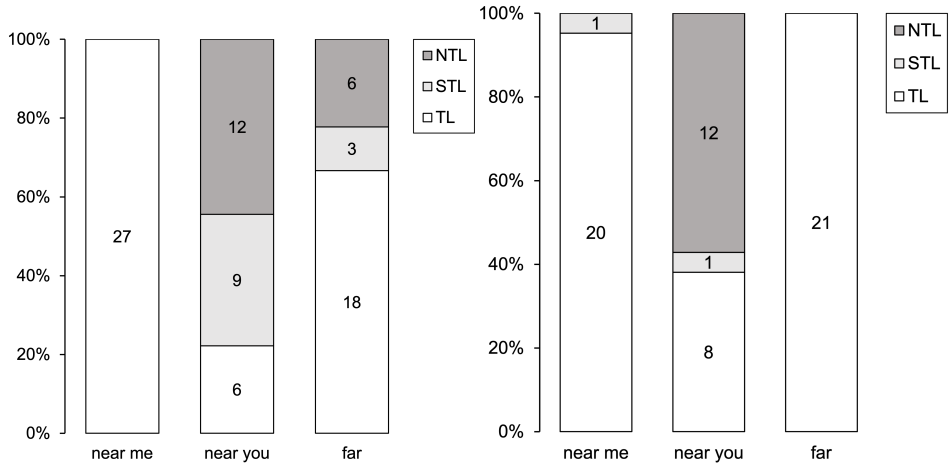


Figure 3: Ternary demonstrative systems: Comprehension and production results (from Terenghi 2022a: 9)

(TL) fashion (that is: compatibly with a three-way deictic opposition) in the overwhelming majority of cases. In hearer-related contexts, instead, both production and comprehension show a considerable amount of non-target-like (NTL) responses, or responses that are not compatible with the hearer-oriented reading. In this context, rather, it can be concluded that the participants perform at chance. This is in line with the predictions made above, once it is assumed that demonstrative systems are syntactically derived by means of person features (Harbour 2016, Bjorkman et al. 2019, Cowper & Hall 2019, Terenghi 2021b, 2023: Ch. 3). Thus, the featural derivation assumed for (6) is given in (7):[3]

(7) a. *queʃtə* (speaker-related deictic domain): +participant(+author(π_χ))
b. *quessə* (hearer-related deictic domain): +participant(−author(π_χ))
c. *quellə* (non-participant-related deictic domain): −participant(−author(π_χ))

Importantly, the conclusion that the non-monotonically derived category alone undergoes loss in heritage speakers was reached by means of the *Microcontact*

[3]Note the the complement set is taken here to be π_χ: this denotes a collection of regions in space, rather than a collection of individuals as π normally does (Harbour 2016), following the discussion in Terenghi (2021b, 2023: 93–94). Also note that the derivation of the non-participant-related domain is yielded by a −participant(−author(...)) sequence: this partly diverges from the discussion in Harbour (2016: 92ff.) and follows Terenghi (2023: 187).

methodology (D'Alessandro 2021, Andriani et al. 2022), whereby heritage languages are considered in different immigration settings: in relation to the phenomenon at hand, this translates into a series of majority languages that display different exophoric demonstrative systems. This was done to assess the role of cross-linguistic influence at a finer-grained level. More precisely, the demonstrative systems of heritage Abruzzese (an upper-southern Italo-Romance variety spoken in a central region of Italy) and Sicilian (an extreme Italo-Romance variety spoken in Sicily) varieties were investigated in contact with Spanish in Argentina, French in Quebec and Belgium and English in the US and in Quebec. Among these, only Argentinian Spanish (and in its prescriptive form) instantiates the same ternary system as that found in the baseline varieties; all other varieties cluster together the hearer- and the non-participant-related domains, yielding a basic two-way opposition between the speaker-related deictic domain (*this near me*) and the non-speaker-related deictic domain (*that far from me*). This latter binary system is found in English and (partly) French, but also in Argentinian Spanish (Kany 1945: 135; Ledgeway & Smith 2016: 888; Andrés Saab, *p.c.*). As shown by Terenghi (2022a), transfer from the different dominant languages is not sufficient to explain these patterns of reduction: the reorganisation of the heritage demonstrative systems does not proceed in a parallel way with respect to that of the relevant dominant language.

3.3 Dual in heritage Arabic varieties

Our second case study focuses on the number domain: we consider the realisation of ternary number systems in heritage Arabic varieties spoken in the US, based on research carried out by Albirini & Benmamoun (2014) and Albirini (2014). The dual number category, which denotes sets of entities with a cardinality of 2, is a feature of classical Arabic but is mostly found as a relic (and typically restricted to body parts that come in pairs) in modern Arabic dialects. However, the Palestinian and Egyptian varieties still display a productive dual category, which is realised by the addition of a dedicated morpheme, *-ein*, to the singular form:

(8) saff → saff-ein 'two classes' (Albirini & Benmamoun 2014: 247)

This contrasts both semantically and morphologically with the plural, which denotes sets of cardinality bigger than 2 and is derived in a non-concatenative fashion (the so-called "broken" plurals):

(9) saff → suffuuf 'classes' (Albirini & Benmamoun 2014: 255)

Albirini & Benmamoun (2014) investigate whether the dual category is still grammaticalised in Palestinian and Egyptian heritage varieties spoken in the US, or whether, possibly under the effect of contact with English, the dedicated dual marker is no longer employed by heritage speakers of these varieties and pairs of entities are referred to analytically (numeral modifier + plural noun, as in English). On the basis of elicited oral production tasks, Albirini & Benmamoun conclude that heritage speakers of Palestinian and Egyptian Arabic in the US are not accurate in forming and using the dual of nouns.

Again, this is in line with our proposal. In fact, in line with the remarks made in Section 3.1, the featural derivation that underlies these different semantics is as follows:

(10) a. *saff* (singular) +minimal(+atomic(P))
b. *saff-ein* (dual) +minimal(−atomic(P))
c. *suffuuf* (plural) −minimal(−atomic(P))

That is, the non-monotonically derived number category is the one that undergoes change and loss in the relevant heritage varieties. Albirini & Benmamoun (2014) and Albirini (2014) suggest that this change might be the effect of transfer from the dominant language, while at the same time highlighting some issues that do not straightforwardly fall out of this. In particular, one of the attested deviant patterns in dual formation is only partly compatible with the English structure: as shown in (11), the deviant realisation of a target dual morphology is analytic, as in English, but the numeral 'two' combines with a singular noun, rather than with the plural one:

(11) Heritage Egyptian Arabic (Albirini 2014: 741)
ʕindi tnein zamiil fi nafs š-ša??a
at-me two roommate in same the-apartment
'I have two roommates in the same apartment.'

This observation cannot conclusively rule out the role of transfer, which can be one of the factors at play in the loss of the dual semantics; future research should examine whether the dual category is unstable, as predicted by the non-monotonicity of the functional sequence that derives it, or not when in contact with comparable ternary number systems. Pending this, it can at least be concluded that heritage Arabic varieties behave in a way that is compatible with the our proposal.

4 Sequences at sentence-level

Complexity at sentential level is more difficult to capture. We will limit ourselves to the case of word order, considering the parametric approach put forward by Roberts (2019), according to which the relevant sequence of features is the one modelled along a parameter hierarchy. The rationale behind this hypothesis is much in the line with the discussion in Section 2: concretely, assuming that syntactic properties are derived by a cluster of parameters relative to the activity of a single feature [F] in differently sized domains, if that feature is not active in the derivation of the relevant phenomenon in a given domain (e.g., all heads), it is absent ([−F]) from it; conversely, if that feature is instead active in a given domain, it is present ([+F]) in it. The domains move from the most general (is the feature active at all?: at the top of the hierarchy) to the most specific one (is the feature active for some specific lexical items only?: at the bottom of the hierarchy). Parametric variation (different parameter settings moving down along the hierarchy) thus derives cross-linguistic variation by means of feature activity along the spine.

Considering word order, recall that word order is determined by one feature, the [EPP], which ensures that a head attracts an XP to its specifier (see e.g. Roberts & Holmberg 2010), as well as head movement:[4] if [EPP] is consistently absent on all heads on the syntactic spine, then the resulting word order will be head-final; if [EPP] is consistently present, then the resulting word order will be head-initial. Non-harmonic word orders are instead derived by an inconsistent setting of the [EPP] parameter: absent in some domains, present in others. This latter configuration is taken to be "marked" and its markedness is in turn brought back to a third-factor principle known as "input generalisation", whereby the learner is taken to generalise the first setting (whether negative or positive) to all subsequent parameters, unless available evidence suggests otherwise (Roberts & Holmberg 2010, a.o.).

Here, we examine two cases of word order change in HL: the first one follows the development of heritage Moundridge Schweitzer German (hereafter MSG), examined by Hopp & Putnam (2015). The second one regards word order in Even and Sakha, two verb-final languages in contact with Russian. Notice that although Even and Sakha are not spoken only in emigrant communities, they have all the features of HLs: they are spoken by minorities, they are heavily exposed to superstratal Russian, and children acquire them as native speakers in a

[4]We use EPP here to refer in general to an "attracting feature", determining movement of either sort: X or XP movement. EPP is, in this sense, more of a generalized diacritic for movement than a proper feature

informal environment (for more information and for the complete data set, the reader is referred to Grenoble & Osipov 2023).

In their 2015 study, Hopp & Putnam show how word order in MSG in contact with English has not moved in the direction of English. MSG is a moribund Palatinate dialect mostly spoken in Kansas, and, like standard German, it presents non-harmonic word order: it is V2 in main clauses and V-final in embedded clauses. Applying Roberts & Holmberg's (2010) Generalization of the input principle, we would expect V2 to be lost in MSG. This is however not the case: V2 remains unscathed, similarly to what is reported for Penn Dutch by Fuller (1997). Recalling Roberts & Holmberg's generalisation in (12a), and assuming that the underlying word order of Germanic languages is OV,[5] we can outline German (and MSG in particular) word order as in (12b):

(12) a. For a class of heads H, *u*EPP for $H_{F:-} \neq v \rightarrow$ $\left\{ \begin{array}{l} [+\text{EPP}]/v_{+\text{EPP}}; \\ [-\text{EPP}] \text{ elsewhere} \end{array} \right\}$

b. In German, *u*EPP for $H \neq v \rightarrow$ $\left\{ \begin{array}{l} [-\text{EPP}] \text{ / } v_{-\text{EPP}}; \\ [+\text{EPP}] \text{ / } C_{root} \end{array} \right\}$

From (12) it is evident that this is a condition of markedness, according to the definition above, as one head bears a different [EPP] value than the rest. This results in a form of non monotonicity, at least in root clauses. We would expect this situation to be "repaired" by HL speakers by the loss of V2. This result would be also in conformity with the ease of processing, as V2 requires an additional movement of the verb into a specific sentence-initial phrase, as well as the filling of its specifier, possibly because of discourse requirements [6]. This prediction is not borne out: Hopp & Putnam (2015) convincingly show that not only V2 is not lost and MSG has not become SVO like English: it is additionally extended to the embedded environment, as shown by (13), where the finite verb *würde* raises across the negation *nicht* to unambiguously reach the second position:

[5]This assumption is not unsubstantiated; German, Dutch and other Germanic languages show head-final characteristics in many environment: numeral, post-positional, as well as adjectival. Furthermore, as shown by many diachronic studies, embedded clauses are more resilient to change and less interested by information structure-related facts. This leads us to conclude that the basic underlying word order in German is head final.

[6]We are assuming here the classical analysis of V2 by den Besten (1977), according to which the verb in V2 constructions moves to C, and its specifier is filled by an XP. We can either say that the [EPP] attracts both the verb to the C head and the XP to its specifier, or that there are two different EPP-diacritics on C, one for the head and one for the specifier.

(13) MSG Participant 122, from Hopp & Putnam 2015: 204
... dass die Verkäuferin *würde* das nicht merken
that the saleslady would that not notice
'that the saleslady would not notice that'

This last piece of information makes the picture more interesting, and begs for some reflection. To start with, while MSG has indeed not turned into an SVO-language, some change has happened nevertheless in the direction of uniformation: the embedded clauses have developed V2. This means that the lack of uniformity in the sequence has indeed been resolved, at least within the verbal domain: C has retained [EPP] in main clauses, but this has been furthermore transmitted to the T/C of the embedded clause, creating a monotonic sequence in the verbal domain.

The second observation is that this change has not started from within, that is: it has not started from an extension of the value of EPP on *v*; rather, it has been induced by some mirroring of the feature value on the C of the matrix clause. This is also not totally unexpected, as change in heritage languages has been argued to penetrate the structure from a peripheral/edge position and slowly extend to the whole sentential domain. Interestingly, this is not the direction followed by first language acquisition, as argued by Roberts & Holmberg 2010: this suggests that HLs have their own mechanisms of adaptation; this reflects the fact that change in contact is less uniform and more idiosyncratic than diachronic change, which has already been noted when discourse elements are involved (see D'Alessandro et al. (forthcoming[a])). Furthermore, this tendency to uniformity and reduction of complexity has taken place within one domain: this is also not unexpected, given that language change in HL takes place step-wise, and optionality and co-existence of different, even conflicting, features is quite common (Polinsky 2018, Aalberse et al. 2019).

Let us now turn to the case of Even and Sakha investigated by Grenoble & Osipov (2023); this is somewhat more complex, given that these minority languages are in contact with Russian, a non-configurational language with a word order which is much less fixed than that of English, for instance. A contact language with a discourse-driven word order is more difficult to generalise upon. For basic declarative clauses, Russian can be considered to be SVO; both Even and Sakha are SOV (Malchukov 1995; Stachowski & Menz 1998). This means that, while Russian is head initial, i.e. according to Roberts has a [−EPP] on *v*, Even and Sakha have [+EPP], being head-final. In a study on word order, Grenoble & Osipov report a shift in younger speakers towards SVO word order. No evidence has been mentioned by the authors about word order change in the language otherwise,

but case has been argued to be undergoing loss. This amounts to saying that the change that Sakha and Even are undergoing is a head-initial to head-final shift in the whole language domain. As an example, consider (14), from Grenoble & Osipov (2023: 33):

(14)

a. S-O1-O2-V	Asatkan	ŋin-u	ulre-č	*ulič-d-de-n*
	girl-NOM	dog-ACC	meat-INST	feed-IPFV-PRS-3SG
b. S-O2-O1-V	Asatkan	ulre-č	ŋin-u	*ulič-d-de-n*
c. S-O1-V-O2	Asatkan	ŋin-u	*ulič-d-de-n*	ulre-č
d. S-V-O1-O2	Asatkan	*ulič-d-de-n*	ŋin-u	ulre-č

'The girl feeds the dog meat.'

Grenoble & Osipov (2023) argue that in traditional Even the only possible word order is V-final. The examples in (14) are instead all acceptable and produced by younger generations of Even speakers. While all word orders co-exist, the mere fact that all four word orders are found in modern Even highlights a change in progress, which they attribute to the contact with Russian.

We leave this here as a speculation, given that sufficient data are lacking, to support this generalisation. What matters is that change, in this context, seems to start off indeed from *v*, like in L1 acquisition contexts and as predicted by Roberts & Holmberg (2010).

Before concluding, it should be noted that our proposal seems to make predictions on the areas of heritage syntax that are more or less vulnerable to transfer. In the case of word order, transfer has been widely indicated as the possible source of change in heritage languages (going beyond the cases discussed here, see discussion in Polinsky 2018: section 6.7 and reference therein). However, it is crucially *not* the case that transfer affects word order across the board.

Heritage Egyptian Arabic varieties illustrate this point quite convincingly. Albirini et al. (2011) show that Egyptian Arabic heritage speakers tend to shift from VSO- to SVO-order in the clausal domain; this is likely to be attributed to transfer from the dominant language, English, that lacks the VSO-order altogether. However, in the nominal domain, Egyptian Arabic heritage speakers consistently retain the "baseline" word order noun–adjective, which is not prone to change under the pressure of English adjective–noun word order (Albirini 2014). Note that the impact of English in this domain is acknowledged in the progressive loss of agreement between the adjective and the noun head. However, transfer does not apply to word order in this case. This fact can be traced back to the general harmony of functional heads already discussed: while a change VSO > SVO does not have an impact on the system, as head-directionality is preserved, a change

NA > AN would amount to a breach of the harmonic setting of parameters across the nominal and the clausal domains. This is expected to be disfavoured, following our proposal. Hence, the heritage Egyptian Arabic data seem to confirm that monotonic sequences of features are favoured in heritage grammars and further constrain the domain of application of transfer: transfer can apply, but only if it does not lead to a cognitively disfavoured sequence of features.

5 Conclusions

The foregoing argued that unrelated changes attested in heritage languages ultimately all hinge on the concept of non-monotonicity along feature sequences: when a feature sequence contains different values, the switching point can be identified as the gateway to change in heritage varieties. Change has been shown to be sometimes further driven by transfer, but transfer itself seems to be to some extent constrained to targeting those switching points along the relevant feature sequences. As such, the ultimate trigger for syntactic change in the domains considered in this study seems to be simply complexity, which is captured straightforwardly by the bias towards monotonic derivations.

One additional observation is that, for the cases illustrated in this work, the predicted change for heritage languages is identical to that which has been documented in diachrony; conversely, no such strict parallel between syntactic change in contact and in diachrony seems to be observable for phenomena which eschew a structural analysis in terms of featural sequences as the ones discussed here. Crucially, the latter phenomena are those about whose development no clear-cut predictions can be made (see D'Alessandro et al. (forthcoming[b])). This seems to suggest that change in contact under a restricted language use situation (heritage languages) and change in diachrony target one and the same structural sequence of syntactic heads (possibly at different speeds, Kupisch & Polinsky 2022), and that otherwise their outputs diverge for more complex, not exclusively syntactic phenomena. While the sensitivity of bilinguals to phenomena sitting at (external) interfaces is well-known (Interface Hypothesis, Sorace & Serratrice 2009, Sorace 2011 for an overview), the relation between this and diachronic change had not been explored before.

In summary, we have shown that languages do indeed undergo a reduction in complexity, but only as far as purely grammatical elements are concerned. Since the same conclusions does not seem to be warranted for syntactic phenomena involving discourse, for which change takes mostly unpredictable paths as determined by several external factors (such as the attitude of the speakers, the context, the level of mastery of the language, *etc.*) as well as by grammar-internal

factors (such as structural similarity), this strongly suggests that a division needs to be drawn between different syntactic phenomena. Therefore, we hope to have shown that it is not methodologically valid to consider complexity of a grammatical system as a whole. Rather, as different subparts of grammatical systems behave differently, depending on whether external factors or other modules are involved or not, it is necessary to suitably delimit the domain of investigation for the issue of complexity in language to be purposefully addressed.

References

Aalberse, Szanne, Ad Backus & Pieter Muysken. 2019. *Heritage languages: A language contact approach.* Amsterdam/Philadelphia, PA: John Benjamins. DOI: 10.1075/sibil.58.

Aikhenvald, Alexandra Y. 2006. Grammars in contact: A cross-linguistic perspective. In Alexandra Y. Aikhenvald & Robert M. W. Dixon (eds.), *Grammars in contact: A cross-linguistic typology*, 1–66. Oxford/New York: Oxford University Press.

Albirini, Abdulkafi. 2014. Toward understanding the variability in the language proficiencies of Arabic heritage speakers. *International Journal of Bilingualism* 18(6). 730–765. DOI: 10.1177/1367006912472404.

Albirini, Abdulkafi & Elabbas Benmamoun. 2014. Aspects of second language transfer in the oral production of Egyptian and Palestinian heritage speakers. *International Journal of Bilingualism* 18(3). 244–273. DOI: 10.1177/1367006912441729.

Albirini, Abdulkafi, Elabbas Benmamoun & Eman Saadah. 2011. Grammatical features of Egyptian and Palestinian Arabic heritage speakers' oral production. *Studies in Second Language Acquisition* 45(2). 273–303.

Andriani, Luigi, Roberta D'Alessandro, Alberto Frasson, Brechje van Osch, Luana Sorgini & Silvia Terenghi. 2022. Adding the microdimension to the study of language change in contact. Three case studies. *Glossa* 7(1). 1–36. DOI: 10.16995/glossa.5748.

Andriani, Luigui & Roberta D'Alessandro. 2022. D,T,C drop in heritage Italo-Romance in NYC: Some descriptive generalizations. In Gabriela Alboiu, Daniela Isac, Alexandru Nicolae, Mihaela Tănase-Dogaru & Alina Tigău (eds.), *A life in linguistics: A festschrift for Alexandra Cornilescu on her 75th birthday*, 41–78. Bucharest: Bucharest University Press.

Barwise, Jon & Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4. 159–219. DOI: 10.1007/BF00350139.

Bayram, Fatih, Diego Pascual y Cabo & Jason Rothman. 2019. Intra-generational attrition: Contributions to heritage speaker competence. In Monica S. Schmid & Barbara Köpke (eds.), *The Oxford handbook of language attrition*, 446–457. Oxford/New York: Oxford University Press.

Biberauer, Theresa. 2017. Factors 2 and 3: A principled approach. *Cambridge Occasional Papers in Linguistics* 10. 38–65. DOI: www.mmll.cam.ac.uk/files/copil_10_3_biberauer.pdf.

Biberauer, Theresa. 2019. Factors 2 and 3: Towards a principled explanation. *Catalan Journal of Linguistics* 2019 Special Issue: Generative Syntax. Questions, Crossroads, and Challenges. 45–88. DOI: 10.5565/rev/catjl.219.

Biberauer, Theresa & Michelle Sheehan (eds.). 2013. *Theoretical approaches to disharmonic word order*. Oxford/New York: Oxford University Press.

Bjorkman, Bronwyn M., Elizabeth Cowper, Daniel Currie Hall & Andrew Peters. 2019. Person and deixis in Heiltsuk pronouns. *Canadian Journal of Linguistics* 64(4). 574–591. DOI: 10.1017/cnj.2019.13.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49. DOI: 10.1016/j.lingua.2012.12.003.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York, NY: Harper & Row.

Clahsen, Harald & Claudia Felser. 2006. Grammatical processing in language learners. *Applied Psycholinguistics* 27. 3–42.

Cowper, Elizabeth & Daniel Currie Hall. 2019. Pronouns and demonstratives in Marshallese: The morphosyntax of person and place. Invited talk presented at the YYC Pronouns Workshop, University of Calgary. https://incl.pl/dch/pdfs/Cowper-Hall-2019-YYC-handout.pdf.

D'Alessandro, Roberta. 2021. Syntactic change in contact: Romance. *Annual Review of Linguistics* 7(1). 309–328. DOI: 10.1146/annurev-linguistics-011619-030311.

D'Alessandro, Roberta, Luigi Andriani, Alberto Frasson, Manuela Pinto, Luana Sorgini & Silvia Terenghi. Forthcoming(a). Heritage languages and syntactic theory. In Roberta D'Alessandro & Michael T. Putnam (eds.), *Microcontact and syntactic theory: A theory of language change in contact, with a special look at microcontact*. Oxford/New York: Oxford University Press.

D'Alessandro, Roberta, Luigi Andriani, Alberto Frasson, Manuela Pinto, Luana Sorgini & Silvia Terenghi. Forthcoming(b). Microcontact and syntactic theory. In Roberta D'Alessandro, Michael T. Putnam & Silvia Terenghi (eds.), *Heritage languages and syntactic theory*. Oxford: Oxford University Press.

D'Alessandro, Roberta, David Natvig & Michael T. Putnam. 2021. Addressing challenges in formal research on moribund heritage languages: A path forward. *Frontiers in Psychology* 12. 700126. DOI: 10.3389/fpsyg.2021.700126.

den Besten, Hans. 1977. On the interaction of root tansformations and lexical deletive rules. In Werner Abraham (ed.), *On the formal syntax of the Westgermania: Papers from the 3rd Groningen Grammar Talks* (Linguistik Aktuell/Linguistics Today 3), 47–132. Amsterdam: John Benjamins.

Diessel, Holger. 1999. *Demonstratives: Form, function, and grammaticalization*. Amsterdam/Philadelphia, PA: John Benjamins. DOI: 10.1075/tsl.42.

Fuller, Janet M. 1997. *Pennsylvania Dutch with a Southern touch: A theoretical model of language contact and change*. Columbia, SC: University of South Carolina. (Doctoral dissertation).

Grenoble, Lenore A & Boris Osipov. 2023. Understanding language shift. *Linguistic Approaches to Bilingualism* 13(1). 122–132. DOI: 10.1075/lab.22100.gre.

Harbour, Daniel. 2008. *Morphosemantic number: From Kiowa noun classes to UG number features*. Dordrecht: Springer. DOI: 10.1007/978-1-4020-5038-1.

Harbour, Daniel. 2011. Descriptive and explanatory markedness. *Morphology* 21. 223–245. DOI: 10.1007/s11525-010-9167-0.

Harbour, Daniel. 2014. Paucity, abundance, and the theory of number. *Language* 90(1). 185–229. DOI: 10.1353/lan.2014.0003.

Harbour, Daniel. 2016. *Impossible persons*. Cambridge, MA: The MIT Press. DOI: 10.7551/mitpress/9780262034739.001.0001.

Haspelmath, Martin. 2006. Against markedness (and what to replace it with). *Journal of Linguistics* 42(1). 25–70. DOI: 10.1017/S0022226705003683.

Hawkins, John A. 1983. *Word order universals*. London: Academic Press.

Hopp, Holger & Michael T. Putnam. 2015. Syntactic restructuring in heritage grammars: Word order variation in Moundridge Schweitzer German. *Linguistic Approaches to Bilingualism* 5(2). 180–213. DOI: 10.1075/lab.5.2.02hop.

Jakobson, Roman. 1932. Zur Struktur des russischen Verbums. In no editor (ed.), *Charisteria Vilhelmo Mathesio oblata*, 74–83. Prague: Cercle Linguistique de Prague.

Jakobson, Roman. 1939. Signe zéro. In *Mélanges de linguistique offerts à Charles Bally*, 143–152. Geneva: Georg et Cie.

Jakobson, Roman & Morris Halle. 1956. *Fundamentals of language*. Berlin/New York: Mouton de Gruyter.

Kany, Charles E. 1945. *American-Spanish Syntax*. Chicago: University of Chicago Press.

Kean, Mary-Louise. 1975. *The theory of markedness in generative grammar*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation). DOI: http://www.ai.mit.edu/projects/dm/theses/kean75.pdf.

Kupisch, Tanja & Maria Polinsky. 2022. Language history on fast forward: Innovations in heritage languages and diachronic change. *Bilingualism: Language and Cognition* 25(1). 1–12. DOI: 10.1017/S1366728921000997.

Laleko, Oksana & Maria Polinsky. 2017. Silence is difficult: On missing elements in bilingual grammars. *Zeitschrift für Sprachwissenschaft* 36(1). 135–163. DOI: 10.1515/zfs-2017-0007.

Ledgeway, Adam & John C. Smith. 2016. Deixis. In Adam Ledgeway & Martin Maiden (eds.), *The Oxford Guide to the Romance Languages*, 879–896. Oxford/New York: Oxford University Press. DOI: 10.1093/acprof:oso/9780199677108.003.0054.

Lyons, John. 1977. *Semantics*, vol. 2. Cambridge, UK: Cambridge University Press. DOI: 10.1017/CBO9780511620614.

Malchukov, Andrej. 1995. *Even*. Munich: Lincom Europa.

McWhorter, John H. 2001. The world's simplest grammars are creole grammars. *Linguistic Typology* 5(2–3). 125–166. DOI: 10.1515/lity.2001.001.

Montrul, Silvina. 2008. *Incomplete acquisition in bilingualism: Re-examining the age factor*. Amsterdam/Philadelphia, PA: John Benjamins. DOI: 10.1075/sibil.39.

Nichols, Johanna. 1992. *Linguistic diversity in space and time*. Chicago/London: University of Chicago Press.

Pires, Acrisio & Jason Rothman. 2009. Disentangling sources of incomplete acquisition: An explanation for competence divergence across heritage grammars. *International Journal of Bilingualism* 13(2). 211–238. DOI: 10.1177/1367006909339806.

Polinsky, Maria. 2018. *Heritage languages and their speakers*. Cambridge, UK: Cambridge University Press. DOI: 10.1017/9781107252349.

Polinsky, Maria. 2022. On how's and why's of heritage languages. Paper presented at Heritage Language Syntax 3, University of Paris.

Polinsky, Maria & Gregory Scontras. 2020. Understanding heritage languages. *Bilingualism: Language and Cognition* 23(1). 4–20. DOI: 10.1017/S1366728919000245.

Roberts, Ian. 2019. *Parameter hierarchies and Universal Grammar*. Oxford/New York: Oxford University Press. DOI: 10.1093/oso/9780198804635.001.0001.

Roberts, Ian & Anders Holmberg. 2010. Introduction: Parameters in minimalist theory. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric variation: Null subjects in Minimalist theory*, 1–57. Cambridge: Cambridge University Press.

Sequeros-Valle, Jose, Bradley Hoot & Jennifer Cabrelli. 2020. Clitic-doubled left dislocation in L2 Spanish: The effect of processing load at the syntax-discourse interface. *Language Acquisition* 27(3). 306–330. DOI: 10.1080/10489223.2020.1769628.

Sheehan, Michelle, Theresa Biberauer, Anders Holmberg & Ian Roberts. 2017. *The Final-over-Final Condition: A syntactic universal*. Cambridge, MA: The MIT Press. DOI: 10.7551/mitpress/8687.001.0001.

Sorace, Antonella. 2011. Pinning down the concept of "interface" in bilinguals. *Linguistic Approaches to Bilingualism* 1(1). 1–33. DOI: 10.1075/lab.1.1.01sor.

Sorace, Antonella & Ludovica Serratrice. 2009. Internal and external interfaces in bilingual language development: Beyond structural overlap. *International Journal of Bilingualism* 13(2). 195–210. DOI: 10.1177/1367006909339810.

Stachowski, Marek & Astrid Menz. 1998. Yakut. In Lars Johanson & Éva Á. Csató (eds.), *The Turkic languages*, 417–433. London: Routledge.

Terenghi, Silvia. 2021a. (Non-)monotonicity effects in the person domain. Talk presented at the 2021 Annual Meeting of the Linguistic Association of Great Britain (LAGB), online, 6–9 September 2021.

Terenghi, Silvia. 2021b. A person-rooted spatial analysis of demonstratives. Talk presented at WCCFL 39, University of Arizona/online, 8–11 April 2021.

Terenghi, Silvia. 2022a. Demonstrative systems are not affected by contact: Evidence from heritage southern Italo-Romance. *Languages* 7(3). 201. DOI: 10.3390/languages7030201.

Terenghi, Silvia. 2022b. Monotonic and non-monotonic person categories: New natural classes. Talk presented at SinFonIJa 15, University of Udine, 22–24 September 2022.

Terenghi, Silvia. 2023. *Missing person: Structure and change in Romance demonstratives*. Utrecht: Utrecht University. (Doctoral dissertation). DOI: 10.48273/LOT0635.

Trubetzkoy, Nicolay. 1931. Die phonologischen Systeme. *Travaux du Cercle Linguistique de Prague* 4. 96–116.

Trubetzkoy, Nicolay. 1939. *Grundzüge der Phonologie*. Göttingen: Vandenhoeck & Ruprecht.

van Osch, Brechje & Petra Sleeman. 2018. Spanish heritage speakers in the Netherlands: Linguistic patterns in the judgment and production of mood. *International Journal of Bilingualism* 22(5). 513–529. DOI: 10.1177/1367006916654365.

Varatharaj, Ashvini, Gregory Scontras & Naomi Nagy. 2024. A multigenerational analysis of heritage language complexity. In Maria Polinsky &

Michael T. Putnam (eds.), *Formal approaches to complexity in heritage languages*, 133–152. Berlin: Language Science Press. DOI: ??.

Formal approaches to complexity in heritage languages

This collective volume breaks new ground in studies of linguistic complexity by addressing this phenomenon in heritage languages. It dismisses with the conception that heritage languages are less complex than their baseline or homeland counterparts and shows complexity trade-offs at various levels of linguistic representation. The authors consider defining properties of complexity as a phenomenon, diagnostics of complexity, and the ways complexity is modeled, measured, or operationalized in language sciences. The chapters showcase several bilingual dyads and offer new empirical data on heritage language production and use.