PREFERENCES IN MINIATURE LANGUAGE LEARNING AS EVIDENCE OF BIASES IN LANGUAGE USE

by

Lucy Jane Kim

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF LINGUISTICS

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2023

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by: Lucy Jane Kim
titled: Preferences in Miniature Language Learning as Evidence of Biases in Language Use

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Janet Nicol — Date: Jul 20, 2023

(Jul 20, 2023 17:58 PDT)
Masha Fedzechkina — Date: Jul 20, 2023

Andrew Wedel (Jul 20, 2023 18:39 PDT)
Andrew Wedel — Date: Jul 20, 2023

Adam Ussishkin (Jul 20, 2023 18:02 PDT)
Adam Ussishkin — Date: Jul 20, 2023

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Janet Nicol — Date: Jul 20, 2023
Linguistics

Acknowledgements

Firstly, I would like to thank my advisor, Masha Fedzechkina, for all of her guidance during my graduate career. Masha, I have learned so much from you and am very grateful to have worked with you. I'd also like to thank my dissertation chair Janet Nicol and my committee members Andy Wedel and Adam Ussishkin. I'm very fortunate to have learned from you as my professors throughout the years and I'm thankful for your support and excellent feedback during the dissertating process.

Next, I would like to thank my colleagues who helped make this dissertation work possible: Florian Hafner and Remo Nitschke for German translations of experiment materials, Remo Nitschke and Masha Alexeeva Zupon for consultation about German and Russian, Jayce Ryrholm for guidance with miniature language learning logistics, and Remo one more time for the LaTeX template I used to write this dissertation. I would also like to thank SBSRI for the Dissertation Research Grant that funded the study in Chapter 4.

I want to thank all of my University of Arizona Linguistics colleagues and friends who have provided me with support and have made my life in Tucson so wonderful. Thank you to my dear friend Yiyun Zhao who constantly helped me grow as a researcher. Thank you to Amber Lubera for the study buddy meetups that helped me finish and the social activities that kept me sane. Thank you to the rest of the best cohort ever: Meg Harvey, Jian Gang Ngui, Othman Alshehri, Kristina Mihajlovic, Kevan Joe, Kerry McCullough, Patricia Lee, Christina Newhall, and Sara Mountain. Thank you to my graduate student mentor Jon Geary for all of his advice and friendship. I also want to thank my other great friends in this department including Jianrong Yu, Seongjin Park, Ryan Walter Smith, Andrew Zupon, Masha Alexeeva Zupon, John W. W. Powell, Remo Nitschke, Megh Krishnaswamy, Luis Irizarry, and Yiseul Han.

Finally, I want to thank my family. Mom and Dad, thank you for showing me how to think creatively and supporting me in whatever I want to do. Peter, thank you for your never-ending love and support and for all of the sacrifices you have made that allowed me to finish my PhD. And also thank you to my dog, Bingsoo, for making me laugh every day while writing my dissertation.

Land Acknowledgment

We respectfully acknowledge the University of Arizona is on the land and territories of Indigenous peoples. Today, Arizona is home to 22 federally recognized tribes, with Tucson being home to the O'odham and the Yaqui. Committed to diversity and inclusion, the University strives to build sustainable relationships with sovereign Native Nations and Indigenous communities through education offerings, partnerships, and community service.

Contents

Contents – *Continued*

Contents – *Continued*

List of Figures

List of Tables

List of Tables – *Continued*

Abstract

Miniature language learning has been widely used to study the causality of language universals. However, factors that have been shown to affect language acquisition in other paradigms are understudied in miniature language learning, which calls into question results that claim to uncover universal biases. In this dissertation, I ask how the production modality of the experiment and the similarity to participants' first language influence miniature language learning behavior to better understand how preferences in miniature language productions are and are not evidence of general cognitive biases underlying language universals.

Chapter 1 introduces the dissertation with a brief review of the concept of language universals and how miniature language learning has been used to study them.

In Chapter 2, I ask whether typing productions rather than speaking productions leads to more regularization of the miniature language. I find no difference in the structure of productions between participants who type and speak their productions, suggesting that either modality can be used for miniature language learning studies.

In Chapter 3, I ask if the structure of participants' first language (L1) affects a well-known miniature language learning result that is claimed to be due to a general cognitive bias. I expose English, German, and Russian speakers to miniature languages and find that speakers of all three languages produce the same type of miniature language output. This finding supports the existence of an L1-independent cognitive bias to balance production effort against message uncertainty.

In Chapter 4, I ask if I would find an L1 influence if the miniature languages were more like participants' L1. I expose German speakers to miniature languages that vary in the degree of direct similarity to German. I find some evidence of L1 influence, but the pattern aligning with the bias to balance production effort against message uncertainty remains for participants learning the miniature lan-

guages closest to German, further supporting the idea that this general cognitive bias is responsible for a statistical language universal.

Chapter 5 summarizes the results of the three studies I presented and discusses how my findings contribute to the discipline of Linguistics.

Chapter 1

Introduction

Linguists and cognitive scientists are interested in the ways in which language is organized and constrained by human cognition. Major strides in uncovering the biases that shape the languages of the world have been made in the study of language universals, including using the experimental paradigm of miniature language learning to show that learners' preferences are aligned with the most crosslinguistically frequent patterns in language. However, in order to claim that miniature language learning behavior is due to general cognitive biases that underlie language universals, researchers must be sure that other biases are not in play. In order to better understand the patterns that we see in miniature language learning and how they do and do not relate to the causality of language universals, this dissertation aims to show how factors aside from general cognitive biases may influence miniature language learning behavior. This chapter will provide the background and context of this research, followed by a brief summary of the objectives and findings of the three studies that make up the body of the dissertation.

1.1 Language Universals

Human languages are incredibly diverse, but the patterns of variation between languages are not wholly random. There are tendencies for certain features and combinations of features to appear in language after language, while other combinations are rarely attested (Greenberg 1963). By way of wide surveys of unrelated languages, it has been revealed that if a language has some property X, it tends to also have property Y (Moravcsik 2011). For example, dominant verb-before-object order languages most commonly place manner adverbs after the verb (e.g., English *run quickly*) while languages with dominant object-before-verb order most commonly place manner adverbs before verbs (e.g., Korean *ppalli* 'quickly' *talliseyyo* 'run') (Dryer 1992). Or consider that in languages with prepositions, the genitive usually follows the governing noun (e.g., Spanish *el libro* 'the book' *de* GEN *la profesora* 'the professor'), while in languages with postpositions, the genitive usually precedes the governing noun (e.g., Hindi *prophesar* 'professor' *kee* GEN *kitaab* 'book') (Greenberg 1963). These *language universals* are of significant interest not only because they reveal how languages pattern, but also because it is theorized that their existence is caused by biases in human cognition (Bever 1970, Fodor 2001). Studying these biases is a key aspect to our understanding of intelligence generally as well as why humans possess the faculty of language (Chomsky 1965).

Typological surveys have done the important work of identifying language universals (Dryer 1992, Hawkins 1983, Greenberg 1963), and in recent years, computational studies using large amounts of corpus data have shown that these language universals align with the most efficient ways to use language (Futrell et al. 2015, Hahn et al. 2021, Kemp et al. 2018). However, these survey methodologies are criticized as it is difficult to find data from enough languages that are not genetically related or in contact with each other for a robust sample (Croft 2017, Dryer 1989, Jaeger et al. 2011). Take the recent example of the correlation between humidity and vowel usage: Everett (2017) claimed by way of typological survey that languages spoken in areas with lower humidity tend to have a smaller number of vowels relative to consonants in their phonemic inventory, but when language family is controlled for, this correlation disappears (Hartmann 2022). Importantly, even in cases where there are enough independent data points, surveys can only show correlation, and not the causation for the patterns found (Fedzechkina et al. 2016). Thus, researchers have turned to other methods to pinpoint the causes for the patterns we see crosslinguistically.

1.1.1 Studying Emergent Languages

One such method is the study of emerging languages. Languages of interest for this line of research are those where there is a severe disruption of language transmission between one generation and the next, thus, the patterns that are identified in the new language are likely to come from general cognitive biases rather than from historic artifacts or language contact (McWhorter 2022, Thomason & Kaufman 1988). These languages include creoles and emergent sign languages.

Creoles, most of which have arisen from European colonization, form when speakers of different languages are brought together (DeGraff 2005). These speakers sometimes first develop a *pidgin* that takes characteristics from the different source languages, but lacks key grammatical features, has multiple forms to express the same meaning, and is the native language of no one. A *creole* emerges when there is a generation of native speakers, and unlike a pidgin, a creole is equal in utility to any other language of the world (Bickerton 2015). Crucially, a creole is distinct from all of its source languages, and it is believed that the speakers introduce grammaticalizations that come from cognitive biases rather than any input language (Bickerton 1984, Michel 1999). Thus, studying what patterns are present across creoles gives more evidence for how cognition shapes language than wider typological surveys, where patterns can be results of historical relationships between languages.

For the same reason, the study of *emergent sign languages* is relevant to the search for the causality of language universals. The most famous of these sign languages is Nicaraguan Sign Language, which began its formation in 1977 when a school for the Deaf brought together Deaf children with little prior language input who began converging on a signed system much like a pidgin (Senghas et al. 1997). As new cohorts entered the school, Nicaraguan Sign Language was transmitted from the older students to the younger, and the signs increased in systematicity, speed, and grammatical complexity from generation to generation of signers, showing a creolization process (Senghas & Coppola 2001). The trajectory of another emergent sign language, Al-Sayyid Bedouin Sign Language, is much the same (Sandler et al. 2011).

McWhorter (2022) summarizes the patterns found across creoles and sign languages. Notably, the author states that (1) emergent languages develop indefinite determiners before definite determiners, (2) all studied emergent languages develop embeddings, and (3) tense markers appear to be secondary to aspect markers. (1) shows that language users have a bias for communicating new information over given information, and (2) and (3) show that recursion and aspect are central properties of human language (McWhorter 2022).

1.2 Miniature Language Learning

To document the trajectory of emergent languages, longitudinal observation of several decades is usually required. In 2005, Hudson Kam & Newport modeled a creolization process over the course of six 30-minute laboratory sessions. The authors employed the experimental paradigm of *miniature language learning*, which involves researchers creating an artificial lexicon and grammar, teaching it to participants over the course of one or a few laboratory sessions, and studying how participants learn or use the lab-created language (Gomez & Gerken 1999, Saffran et al. 1996). Hudson Kam & Newport (2005) created a miniature language with variable input: for 60% of sentences chosen at random, a determiner would precede the noun while the other 40% of sentences did not contain determiners. This probabilistic input mimics a pidgin-creole situation where language learners are exposed to multiple forms with the same meaning. When participants produced their own sentences in the miniature language, the researchers found that adults matched the proportion of determiner use that they were exposed to, but that children (mean age six years) regularized the input. That is, when exposed to an input proportion of 60% of sentences with determiners, children either produced all sentences with determiners or all sentences without determiners at test. This result gave experimental support to the theory that children are arbiters of language change who drive the grammaticalization process of creoles, as children's bias to regularize would eliminate inconsistent forms found in a pidgin (Bickerton 1984). Hudson Kam & Newport's study also led to a proliferation of studies using the technique of exposing participants to variable input and studying how they change the miniature language. What has been found is that depending on the properties of the miniature language, adults do not always match the input. Instead, they often change the input distribution in ways that mirror language universals.

For example, consider this language universal described by Greenberg (1963): Many more languages put the adjective and numeral on the same side of the noun (e.g., English *green books*; *six books*) than languages that put the adjective and numeral on opposite sides of the noun (e.g., French *livres* 'books' *verts* 'green'; *six* 'six' *livres* 'books'). Culbertson et al. (2012) set out to test the causality of this language universal with miniature languages where the position of the adjective and numeral varied probabilistically with respect to the noun. What they found was that at test, adult participants preferred orders where the adjective and numeral were on the same side of the noun, lining up with the language universal. Culbertson et al. posit that this result is due to a cognitive bias that makes symmetric patterns (e.g., noun-before-adjective and noun-before-numeral) easier to learn than asymmetric patterns (e.g., adjective-before-noun and noun-before-numeral).

Or consider another crosslinguistic pattern: Languages with more fixed constituent order (e.g., English, Irish, Swedish, Italian) tend not to mark case while languages with more flexible constituent order (e.g., German, Russian, Japanese, Tamil) tend to mark case (Blake 2001, Futrell et al. 2015, Sapir 1921). Fedzechkina et al. (2017) found that miniature language learners mirrored this crosslinguistic trade-off between constituent order flexibility and case marking. Fedzechkina et al. created two miniature languages, one with fixed constituent order and one with flexible constituent order. Both miniature languages had the object noun case-marked in two-thirds of the input sentences. The researchers found that participants who learned the fixed order miniature language used less case in their productions than the participants who learned the flexible order miniature language. Fedzechkina et al. explain this result in terms of another cognitive bias, one for balancing production effort (as the case marker takes extra effort to add to the sentence) and message uncertainty (as who is doing what to whom is more uncertain in a language with flexible constituent order). This bias is more fully explained in Chapter 3.

Many other studies have found that miniature language learners' outputs mirror the most crosslinguistically frequent patterns, including those patterns involved in phonology (Newport & Aslin 2004, Finley & Badecker 2008, Wilson 2006), semantics (Saratsli et al. 2020), word order (Christiansen 2000, Fedzechkina et al. 2018, Saldana et al. 2021), morphology (Fedzechkina et al. 2012, Kurumada & Grimm 2019, Mansfield et al. 2022, Sheehan et al. 2019), and word length (Kanwal et al. 2017). However, in order to claim that these results are caused by general cognitive biases including the bias for learning symmetric patterns and the bias to balance production effort against message uncertainty mentioned above, researchers must be sure that other biases are not affecting participants' miniature language learning preferences. This dissertation looks at two understudied factors that could influence miniature language learning outcomes.

1.2.1 Factor 1: Production Modality

As it is a behavioral paradigm, miniature language learning is not immune to task effects. For example, it has been found that when learners are asked to produce sentences in a miniature language, they change the input language more when the task has increased memory demands (Hudson Kam & Newport 2009, Hudson Kam & Chang 2009) as well as when they are told that the goal of the task is to align with another speaker rather than to use the miniature language accurately (Perfors 2016). Another factor that may influence miniature language learning outcomes is the modality of participants' productions. The production phase of a miniature language learning study involves participants using the miniature language that they have learned, but studies differ in whether they have participants speak their productions (e.g., Fedzechkina et al. 2012) or whether participants type their productions in a text box (e.g., Smith & Wonnacott 2010). There is some reason to believe that speaking and typing may lead to different results. For one, on statistical learning tasks, where participants discern patterns based on their frequencies in a distribution, participants are better able to track probabilities when they are presented with a distribution aurally than visually (Saffran 2002). In miniature language learning tasks used to investigate the causality of language universals, the input language often has elements that are probabilistically present in order to leave room for participants to induce changes into the language (Fedzechkina et al. 2016). Having a better or worse ability to track this probability distribution could lead to different rates of regularization of a miniature language (Hudson Kam & Newport 2009). Additionally, by speaking their productions, participants in experiments with this modality may be getting extra feedback as they hear themselves say their answer, while participants typing their productions may not have the same benefit (Baese-Berk & Goldrick 2009). There is thus some reason to believe that production modality could affect the acquisition and use of a miniature language. By treating the two modalities as equal without an empirical test, miniature language learning researchers may be attributing participants' behavior to general cognitive biases when the real cause is a task effect.

1.2.2 Factor 2: L1 Influence

Moreover, the language(s) that participants already speak may influence how they learn and use a miniature language. Miniature language learning can be considered a special case of second language (L2) learning where a small linguistic system is acquired in a short period of time. This position is supported by the work of Ettlinger et al. (2016), who found that L2 learners who performed better on classroom learning of Spanish also performed better in miniature language learning tasks. Concerns about the possibility of first language (L1) influence on miniature language learning tasks have been raised often in the field (Culbertson et al. 2012, Fedzechkina et al. 2018, Goldberg 2013). Typically, researchers attempt to avoid L1 influence by making miniature languages as different from participants' L1s as possible, ideally by testing structures that are not present in participants' L1 (e.g., Fedzechkina et al. 2012 looking at optional case marking by testing speakers of English, a language without optional case marking) (Fedzechkina et al. 2016). However, there is no agreed upon measure of how different participants' L1 should be from the miniature language to not induce L1 influence. This is a difficult question to answer as the literature on L1 influence even in the acquisition of a natural L2 or L3 is mixed

(Bardel & Falk 2007, Flege 1989, McAllister et al. 2002, Rothman & Cabrelli-Amaro 2010, Schepens et al. 2015, Slabakova 2017), but in the case of miniature language learning, there have been few investigations into what affect participants' L1 has on results that are attributed to cognitive biases that underlie language universals, and the results of these investigations are ambiguous (Culbertson et al. 2020, Martin & Culbertson 2020).

1.3 The Dissertation

In this dissertation, I look at the factors of production modality and L1 influence and how they may affect miniature language learning behavior that has previously been attributed to general cognitive biases underlying language universals. By understanding the role that these factors play in a miniature language learning experiment, researchers can be more confident about what truly causes miniature language learning behavior, resulting in a methodology that can more precisely probe the causality of language universals. Over the course of three studies, my results show support for the presence of general cognitive biases in language use and provide guidance for best practices in miniature language learning research.

1.3.1 Structure of the Dissertation

The body of this dissertation consists of three articles, with each article reporting on a a miniature language learning study. The articles build on each other in the following way: Chapter 2 gives the justification for the methodology in Chapters 3 and 4, and Chapter 3 provides the research question for Chapter 4. Chapter 5 concludes the dissertation with a summary of my findings and how they fit into the larger literature of Linguistics.

1.3.2 Brief Summary of Body Chapters

Chapter 2

Here, I explore the effect of production modality on miniature language learning behavior. I expose participants to a miniature language with flexible constituent order and measure if the amount that they regularize the language's constituent order depends on whether they type or speak their productions. I hypothesize that typing productions leads to poorer learning outcomes which would cause more regularization of the miniature language. However, I find no difference in regularization behavior between participants who speak and type their productions at test. Because of the practical advantages of collecting typed responses, I conclude by recommending a typing production modality for miniature language learning studies.

Chapter 3

I ask if the structure of participants' L1 affects a well-known miniature language learning result that is claimed to be due to a general cognitive bias to balance production effort against message uncertainty. I expose English (fixed constituent order, no case marking on nouns [1]), German (somewhat flexible constituent order, case system), and Russian (flexible constituent order, rich case system) speakers to miniature languages with either fixed or flexible constituent order and optional case marking. The flexible order miniature language has abstract similarity to German and Russian, but I find that speakers of all three languages produce the same type of miniature language output: they match the input constituent order proportions and use more case marking in the flexible order language than in the fixed, thus following a bias to balance production effort against message uncertainty. These results suggest that the bias to balance production effort against message uncertainty is independent of the structure of miniature language learners' L1.

[1] English makes a nominative/accusative distinction for first and third person animate pronouns (I/me, he/him, she/her, we/us, they/them) but case marking is absent from nouns (Leech 1975)

Chapter 4

Here, I ask a follow-up question to Chapter 3: Would I find an L1 influence if the miniature languages were more similar to participants' L1? I expose German speakers to miniature languages with fixed or flexible constituent order and optional case marking that vary in the degree of direct similarity to German constituent order and case marking. I hypothesize that as the miniature languages get closer to German, German speakers will abandon the bias to balance production effort against message uncertainty in favor of producing more German-like sentences. I find mixed results as German speakers followed the bias to balance production effort against message uncertainty in the miniature languages that had non-German-like constituent ordering with non-German-like case marking and German-like constituent ordering with German-like case marking, but not in the miniature language with non-German-like constituent ordering with German-like case marking. I discuss possible explanations for this U-shaped effect and conclude by recommending that researchers continue to construct miniature languages with the goal of being dissimilar to the languages that participants speak to avoid the introduction of L1 influences.

Chapter 2

Influence of Production Modality on Constituent Order Use

2.1 Introduction

The miniature language learning experiments used to study biases in language use typically have three types of blocks: exposure blocks where participants are presented with miniature language words and phrases paired with pictures or videos representing their meaning, comprehension blocks where researchers can validate that participants understand the meaning of miniature language elements, and production blocks, where participants produce their own utterances in the miniature language and researchers measure how participants actually use the language (Culbertson et al. 2012, Fedzechkina et al. 2017, Smith & Wonnacott 2010). Productions take one of two forms: written, where participants use their computer keyboards to type words and sentences, or spoken, where participants say words and sentences into a microphone or to an experimenter. Using spoken productions in a miniature language learning task can be advantageous when working with pre-literate children, and to simulate a real language use environment as language is primarily spoken. However, when working with adults, there are an increasing number of methodological advantages for using typed productions. One key advantage is in the ability to complete studies over the web. Crowdsourcing platforms offer access to participants that researchers are unable to reach in undergraduate university student populations, resulting in a more diverse subject pool (Casler et al. 2013). However, not all crowdsourcing platform participants have access to a microphone, and among those who do, recording quality can be too low for the researcher to clearly make out the recorded productions. Thus, eliciting typed productions is more practical when recruiting participants who will be using their personal equipment to complete the study. Another advantage is in analysis time. Experimenters can easily write scripts to code written productions for analysis, but spoken productions in an artificial language need to be listened to by a human ear and transcribed before they can be analyzed. Additionally, as texting and IMing become increasingly commonplace forms of communication, the concern over naturalness of typed productions is diminishing. Several researchers have called for using SMS format as an instruction modality for L2 learners since 2009, citing texting as an essential communication method (Li & Cummins 2019). Typing as a miniature language learning modality thus offers practical benefits and can still model communicative behavior. However, using a different production modality may lead to different results. To date, there has been no direct comparison between paradigms that use a spoken production modality versus a typed modality. If we wish to draw comparisons between studies previously run using a spoken modality and new studies using a typed modality, we need to be sure that participants are learning and using the miniature languages the same way.

One concern is that a typed production modality may not provide the same learning benefits as a spoken production modality, which could mean that participants who type their productions in miniature language learning studies do not learn the miniature language as well. In learning non-linguistic patterns that rely on the same probability tracking mechanism as miniature language learning tasks, participants show better performance when exposed to auditory stimuli than when exposed to visual stimuli (Saffran 2002, Conway & Christiansen 2005). This effect is attributed to the sequential nature of auditory processing; the fact that sounds are perceived over time aids participants in task that requires them to track a sequence, while visual stimuli, which can be perceived simultaneously, do not offer this benefit (Saffran 2002). When considering production specifically, there is additional evidence that spoken productions may offer learning benefits over typed productions. In speech processing, there is evidence that speakers engage in a feedback mechanism in which they make future articulation decisions based off of what they have perceived from their own speech (Baese-Berk & Goldrick 2009). Thus, it is possible that participants who voice their productions may have an additional feedback avenue over participants who type their productions, especially if those who type their productions do not reread what they have typed. The additional feedback offered in the spoken modality may lead to better learning. If typing productions leads to participants not learning the miniature language as well as speaking productions, this could result in excluding more participants from analysis for failure to learn the language, which is a costly and time-consuming process, or in an analysis of language use that is not comparable to previous studies because participants do not have the same mastery of the miniature language.

There is a need to systematically test whether typing productions results in a difference in performance in miniature language learning experiments. We take up this question by looking at how participants use constituent ordering in a miniature language that has variable constituent ordering in the input. When variability exists in language input, language users can either reproduce the variability or reduce it through a process called *regularization*. In miniature language learning studies where participants learn a language with unpredictable variation, children often regularize this variation by choosing to use one variant over the other(s) in the production phase of the experiment (Hudson Kam & Newport 2009, 2005, Austin et al. 2022). One explanation for the regularization behavior is that children's limited cognitive capacities lead them to have difficulty tracking variability. Indeed, when adults, who tend to more faithfully reproduce variability in miniature language learning experiments, are given increased cognitive demands, they regularize variable distributions as well (Hudson Kam & Newport 2009, Perfors 2012, Hudson Kam 2019, Hudson Kam & Chang 2009, Ferdinand et al. 2019). For example, Hudson Kam and Chang (2009) found that adult learners regularized a miniature language more when they had to remember lexical items during the production phase than if they were given a word bank. If the typing modality decreases the ability to track probabilities or a feedback mechanism important for learning as compared to the speaking modality, we would expect a similar effect where participants who type their productions have increased cognitive demands because they have learned the language less well. If this is the case, participants who type their productions should regularize more than participants who speak their productions. Here, we expose adult participants to a miniature language with variable constituent order and ask whether regularization of this variability in the output produced depends on whether participants type or speak their productions.

2.2 Methods

2.2.1 Language

The miniature language consisted of simple transitive sentences that were either VSO (67%) or VOS (33%). Constituent order was not conditioned on any aspect of the input. Each noun and verb occurred in both VSO and VOS sentences. Each noun occurred both as the subject and the object of sentences.

The artificial lexicon consisted of ten words (six nouns and four verbs). All words were phonotactically legal non-words of English (Table 2.1). Individual words were synthesized using the Mac Speech Synthesizer in the Greek-accented Melina voice [1] and concatenated into sentences using Praat with 35 ms of silence between words. To ensure there were no artificial associations between a word form and constituent order, two lists were made in which the words were rotated to different meanings and ten participants in each condition were assigned to each list.

2.2.2 Production Modalities

Participants were randomly assigned to complete the experiment in one of two conditions: typing or speaking. The typing condition required participants to type miniature language output into a textbox while the speaking condition required participants to click on a button and record themselves saying their miniature language productions out loud. The only other difference between the speaking and typing conditions was that during learning blocks, participants in the typing con-

[1] A Greek accent was used as this study was conducted to investigate what production modality should be used in Chapter 3, in which speakers of English, German, and Russian participate so the accent needed to be unlike any of the participants' first language.

Table 2.1: Vocabulary lists used in experiment

	Nouns		Verbs	
	Word	Meaning	Word	Meaning
List A	pilka	mountie	tegat	knee
	balti	chef	felmik	tap
	kusa	ref	dopsan	punch
	nettli	bandit	sunid	hug
	flaki	conductor		
	mitsa	hunter		
List B	kusa	mountie	felmik	knee
	mitsa	chef	sunid	tap
	balti	ref	tegat	punch
	pilka	bandit	dopsan	hug
	nettli	conductor		
	flaki	hunter		

dition were presented with miniature language words and sentences both auditorily and written on the screen, while participants in the speaking condition were only presented with miniature language input auditorily. Previous experiments that used a similar procedure to ours had participants speak their productions and only presented the miniature language input auditorily (Fedzechkina et al. 2017, 2018), so we did the same in our speaking condition. There was a need to introduce written forms for the typing condition because without exposure to the orthography, typed productions would not be consistent across participants and the criteria for lexical mistakes in analysis (see Section 2.2.5) would be ambiguous.

2.2.3 Participants

Native English-speaking participants living in the United States were recruited via Amazon Mechanical Turk. Following Fedzechkina et al. (2018), recruitment continued until 20 participants were included in each condition (typing, speaking). 58 participants were recruited in total. Participants were excluded for not speaking English as their L1 (one in typing condition; native language Bikol), using non-input constituent orders on more than 30% of test trials (two in speaking condition; both used SVO order), and producing utterances that could not be coded for constituent order on more than 30% of test trials (three in typing condition, twelve in speaking condition; coding described in Section 2.2.5). This resulted in the final sample of 40 participants total. The experiment took participants an average of 44 minutes to complete. Each participant was compensated $5.

2.2.4 Procedure

Participants were instructed that they would be learning an alien language from a character named Lumi. The experiment was broken up into 11 blocks focusing on different aspects of learning (listed below) aimed at teaching participants first nouns in isolation, then sentences (Figures 2.1- 2.2). The experiment was presented via the online presentation platform FindingFive (FindingFive Team 2019). Language stimuli were presented auditorily and visually (using English-like spelling) in the typing condition and auditorily alone in the speaking condition.

Noun Blocks

The experiment began with participants learning the nouns in isolation. The sequence of Noun Exposure, Noun Comprehension, and Noun Production was presented twice before moving onto the sentence blocks.

Noun Exposure Participants were shown pictures of each of the 6 characters paired with the characters' names. In both conditions, participants were instructed to listen to how Lumi would say the characters' names and repeat the names out loud. Participants could click to replay the audio as many times as they liked. (Noun Exposure 1: twelve trials; Noun Exposure 2: twelve trials).

Noun Comprehension Participants were presented with a name (by audio and text in the typing condition, by audio alone in the speaking condition) and had to pick the correct character from a choice of 4. Participants could click to replay the audio as many times as they liked. Participants received feedback indicating whether

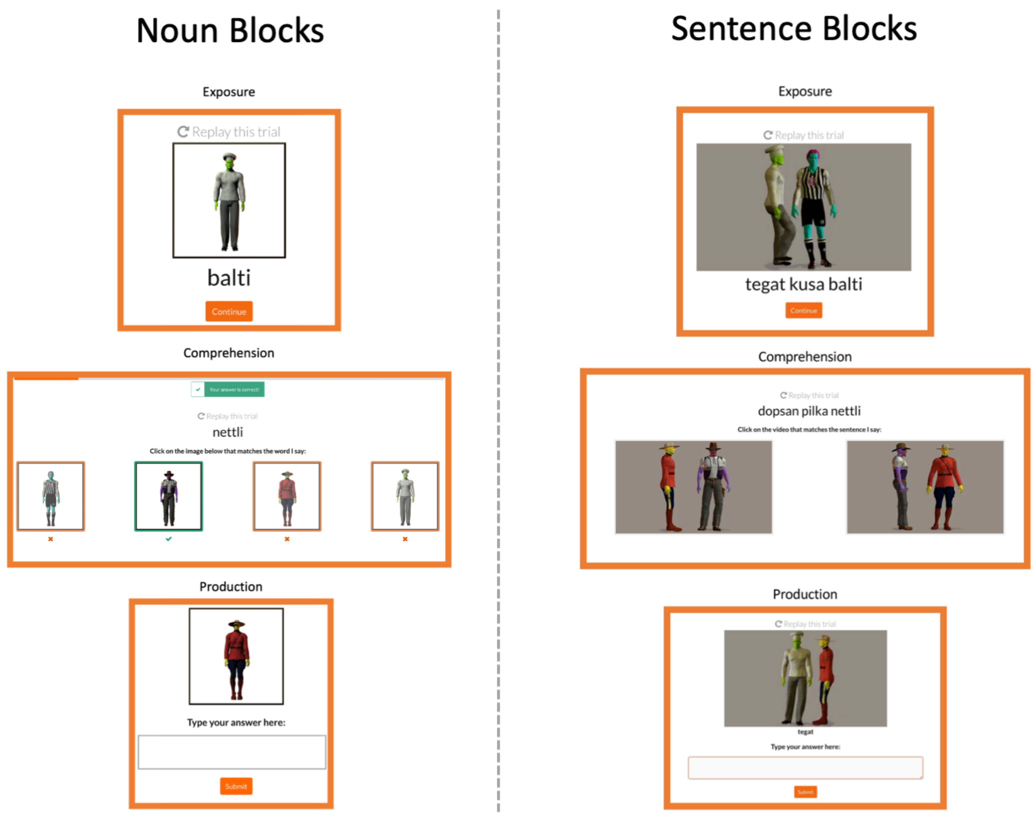

Figure 2.1: Screenshots from all blocks of the typing condition. The speaking condition was identical except that there were no written miniature language words on the screen and participants had a button to record their productions rather than a textbox to type them.

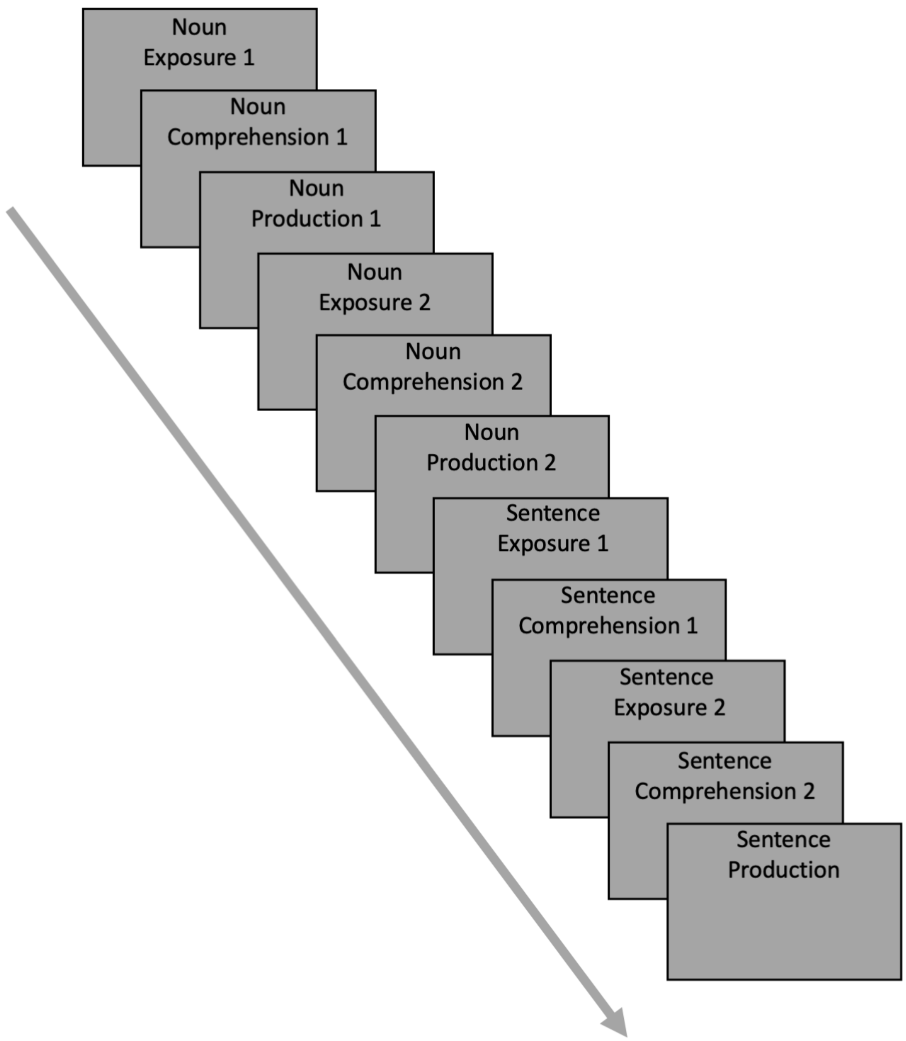

Figure 2.2: Overview of procedure blocks

or not they had chosen the correct character (Noun Comprehension 1: twelve trials; Noun Comprehension 2: twelve trials).

Noun Production Participants were presented with a picture of a character and typed/said the name. They received feedback with the correct name of the character (Noun Production 1: six trials; Noun Production 2: twelve trials).

Sentence Blocks

In sentence blocks, participants were exposed to a sentence exposure followed by a sentence comprehension block twice, before engaging in the critical test, sentence production.

Sentence Exposure Sentence exposure blocks consisted of simple transitive sentences paired with a video of one character doing one action to another character. Participants could click to replay the audio and video as many times as they liked. (Sentence Exposure 1: 48 trials; Sentence Exposure 2: 48 trials).

Sentence Comprehension Participants were provided with a sentence (by audio and text in the typing condition and by audio alone in the speaking condition) and had to select between two videos that differed in which character was the subject and which was the object of the same action. Participants could click to replay the audio and video as many times as they liked. No feedback was provided (Sentence Comprehension 1: 24 trials; Sentence Comprehension 2: 24 trials).

Sentence Production Participants were given previously unseen videos of 2 characters participating in a transitive action and had to type/say a sentence in the language that described who was doing what to whom. Participants were provided with a verb prompt (by audio and text in the typing condition and by audio alone in the speaking condition) to facilitate production. Participants could click to replay the audio and video as many times as they liked. No feedback was provided (Sentence Production: 24 trials).

2.2.5 Coding

For the speaking condition, audio productions were first transcribed using CMU Pronouncing Dictionary form (Han 2013). Then, for both speaking and typing conditions, we assessed lexical accuracy using soft string matching with Levenshtein distance. Levenshtein distance measures the numbers of deletions, insertions, and substitutions between two words. Any word where the Levenshtein distance between the target form and the production was more than two was coded as a lexical error. For example, the production *metsu* for target *mitsa* was coded as a lexical error because it involves two substitutions, but the production *mits* was not coded as a lexical error because it involves only one deletion. We then coded sentence comprehension selections and sentence production utterances for constituent order. If a participant made lexical errors on both nouns, constituent order could not be determined for that trial. These trials were labeled as uncodeable. If a participant did not use two nouns and one verb in their utterance, constituent order could not be determined for that trial. These trials were also labeled as uncodeable. Some participants produced sentences with constituent orders other than the VSO and VOS present in the input (most common other constituent order was SVO). We coded non-VSO or non-VOS productions as grammatical errors (Table 2.2). Participants with uncodeable utterances or grammatical errors on more than 30% of sentence production trials were excluded from analysis. For remaining participants, trials with uncodeable utterances or grammatical errors were excluded from analysis (6.4% of total sentence production trials).

Table 2.2: Coding Examples

Example Production	Lex. Error	Uncodeable	Gramm. Error	Explanation
`tegat mitsa balti`	N	N	N	This production has no lexical errors and is in VSO order.
`tegat mits balti`	N	N	N	‘mits’ is only a Levenshtein distance of 1 from ‘mitsa’ so is not counted as a lexical error.
`tegat metsu balti`	Y	N	N	‘metsu’ is a Levenshtein distance of 2 away from all lexical items and is counted as a lexical error. Constituent order can be determined from the other two words in the production.
`tegat metsul babi`	Y	Y	N/A	‘metsul’ and ‘babi’ each have a Levenshtein distance over 2 from all lexical items. Constituent order cannot be determined because we cannot identify the subject and object.
`mitsa balti`	N	Y	N/A	There is no verb in this production, so constituent order cannot be determined.
`mitsa tegat balti`	N	N	Y	This production is in SVO order, which is not present in the miniature language.

2.3 Results

2.3.1 Noun Learning Accuracy

We first assess the accuracy of noun learning in the miniature language by looking at participants' selections in the noun comprehension blocks where they selected an image that matched the presented noun and the noun production blocks where they produced the noun that went with the presented image. For participants included in analysis, we calculated the mean percent of trials with errors by production modality condition for each block (Table 2.3). In each block, participants in the speaking modality condition made more errors than participants in the typing modality condition. The difference in mean errors between the modality conditions is larger in the production blocks than the comprehension blocks, evidenced by the fact that the 95% confidence intervals between the conditions do not overlap for the production blocks. The error rate for participants in the speaking modality condition remained high in the last block of noun learning, Production 2 (Mean error rate of 25.0%). Though the confidence intervals do overlap, it is interesting that in Noun Comprehension 1, a block that occurred before any production (Figure 2.2), participants in the speaking condition make more errors than participants in the typing condition. At that point in the experiment, the only difference between conditions was the accompanying written lexical item present in the typing condition but not the speaking condition. This could suggest that the higher error rate we see for participants in the speaking condition is due to the absence of additional written exposure.

2.3.2 Constituent Order in Sentence Comprehension

In the sentence comprehension blocks, participants were presented with a sentence and chose between videos that had a VSO and a VOS interpretation. If typing productions negatively affects learning of the miniature language, we would expect that participants who type their productions would have a harder time tracking the input language distribution, and might overselect one constituent order as compared

Table 2.3: Errors in noun learning blocks

Block	Condition	Percent trials with errors Mean	95% CI
Comprehension 1	Typing	18.8%	9.2-29.2%
	Speaking	30.3%	19.3-41.7%
Comprehension 2	Typing	5.0%	1.3-9.6%
	Speaking	5.8%	2.1-11.3%
Production 1	Typing	29.2%	16.7-42.5%
	Speaking	58.3%	44.2-71.7%
Production 2	Typing	7.9%	2.1-15.4%
	Speaking	25.0%	17.1-32.9%

to participants who speak their productions. We calculated the mean percentage of trials where participants selected a VSO interpretation (Figure 2.3). In both sentence comprehension blocks, participants in the typing condition selected more VSO interpretations than participants in the speaking condition (73.1% vs. 57.7% in Block 1 and 78.1% vs. 65.2% in Block 2). However, in both blocks, the 95% confidence intervals for the two conditions overlap, so there is not strong evidence of a difference in selection behavior between the typing and speaking conditions. Further, the input language 67% VSO is within the 95% confidence intervals for both conditions in both blocks, suggesting that participants did not overselect one constituent order as compared to the input distribution in either production modality condition.

2.3.3 Regularization in Sentence Production

To assess if one production modality leads to stronger regularization behavior than another, we look at the constituent order of the sentences participants produced at our sentence production test block. In Hudson Kam & Newport (2009), regularization was measured as the amount of the dominant variant used, which in our case, would be amount VSO used as our input language had the majority of sentences in VSO order. However, when looking at the constituent order proportions of participants' productions, we found that while some participants regularized by using more

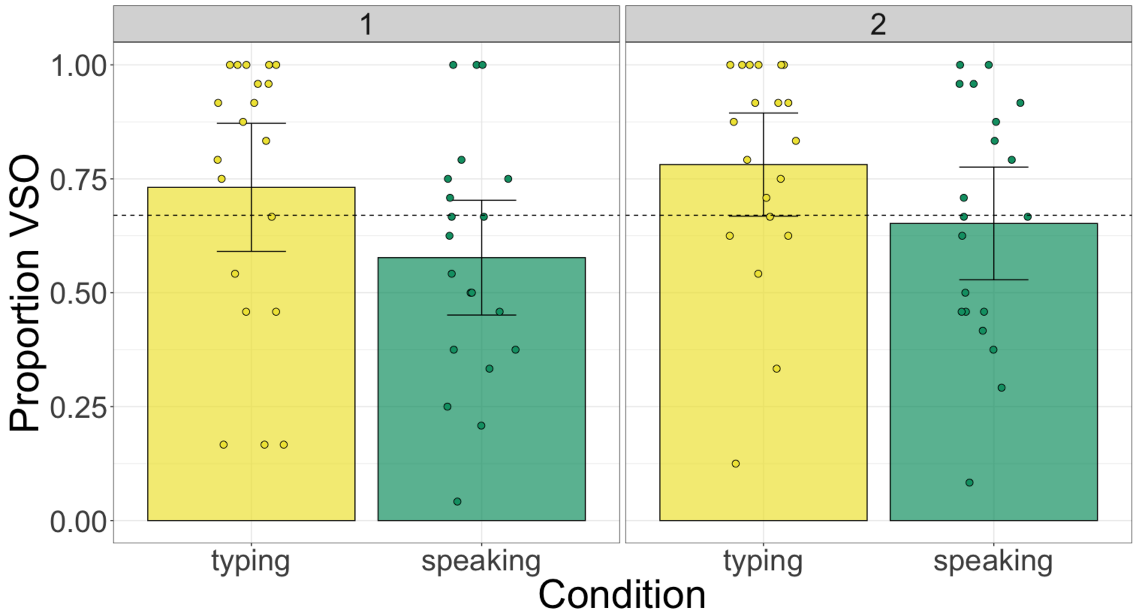

Figure 2.3: Proportion VSO selected during Sentence Comprehension Block 1 (left) and Block 2 (right). The dashed line represents the 0.67 input language proportion. Dots represent individual participants' proportion VSO selection. Error bars are bootstrapped 95% confidence intervals.

VSO than the input language, other participants reduced variability in the language by using more VOS constituent order (Figure B.1). Ferdinand et al. (2019) measure regularization in a different way, by looking at the entropy of participants' output languages. Entropy measures the uncertainty, or variability, in a probability distribution (Shannon 1948). If participants engage in more regularization, or in other words reduce variability in the language, their output language should have lower entropy than the input language. Entropy is calculated with the equation

$$H(V) = -\sum_{v_i \in V} p(v_i) log_2 p(v_i) \tag{2.1}$$

where V is the set of variants, $p(V)$ is the probability distribution of the variants, and $p(v_i)$ is the probability of the ith variant in the set of variants (Ferdinand et al. 2019). In our case, V is the set of constituent order variants and $p(v_i)$ corresponds to the proportion of sentences in the ith constituent order. The entropy of our input language where two-thirds of sentences are in VSO order and one-third of sentences are in VOS order can thus be calculated as

$$H(V) = -\frac{2}{3}log_2\frac{2}{3} + -\frac{1}{3}log_2\frac{1}{3} = 0.918 \text{ bits} \quad (2.2)$$

Here we are interested in whether the entropy in the output languages of participants who type their productions is lower than that of participants who speak their productions. We calculated the entropy for each participant's output language (Figure 2.4) and fit a linear regression that predicts entropy from condition (dummy coded; speaking = 0, typing = 1). In sentence production, we find that participants who type their productions do not have significantly different entropy in their output language than participants who speak their productions ($\hat{\beta} = -0.019, t = -0.129, p = 0.898$; full results in Table A.1). Therefore, we do not find evidence that participants who produce their sentences in the typing modality regularize more than participants who produce their sentences in the speaking modality at our critical test.

The participants who typed their productions did not regularize significantly more than participants who spoke their productions, but both groups did regularize the input language. To measure this behavior, we fit a linear regression model that predicts entropy from production modality condition (dummy coded) with an offset of 0.918 corresponding to the input entropy. The intercept of this model reveals whether the reference level condition differs significantly from the input language entropy. We ran this model twice, once with each condition as the reference level. We found that participants in the typing condition produced languages with significantly lower entropy than the input (Mean entropy 0.497 bits; $\hat{\beta} = -0.421, t = -4.086, p > 0.001$; full results in Table A.2) and that participants in the speaking condition produced languages with significantly lower entropy than the input (Mean entropy 0.516 bits; $\hat{\beta} = -0.402, t = -3.903, p > 0.001$; full results in Table A.3). This shows that participants in both conditions significantly regularized the miniature language constituent order, but as there was no significant difference in entropy between the conditions, we do not find evidence that this regularization behavior is due to production modality.

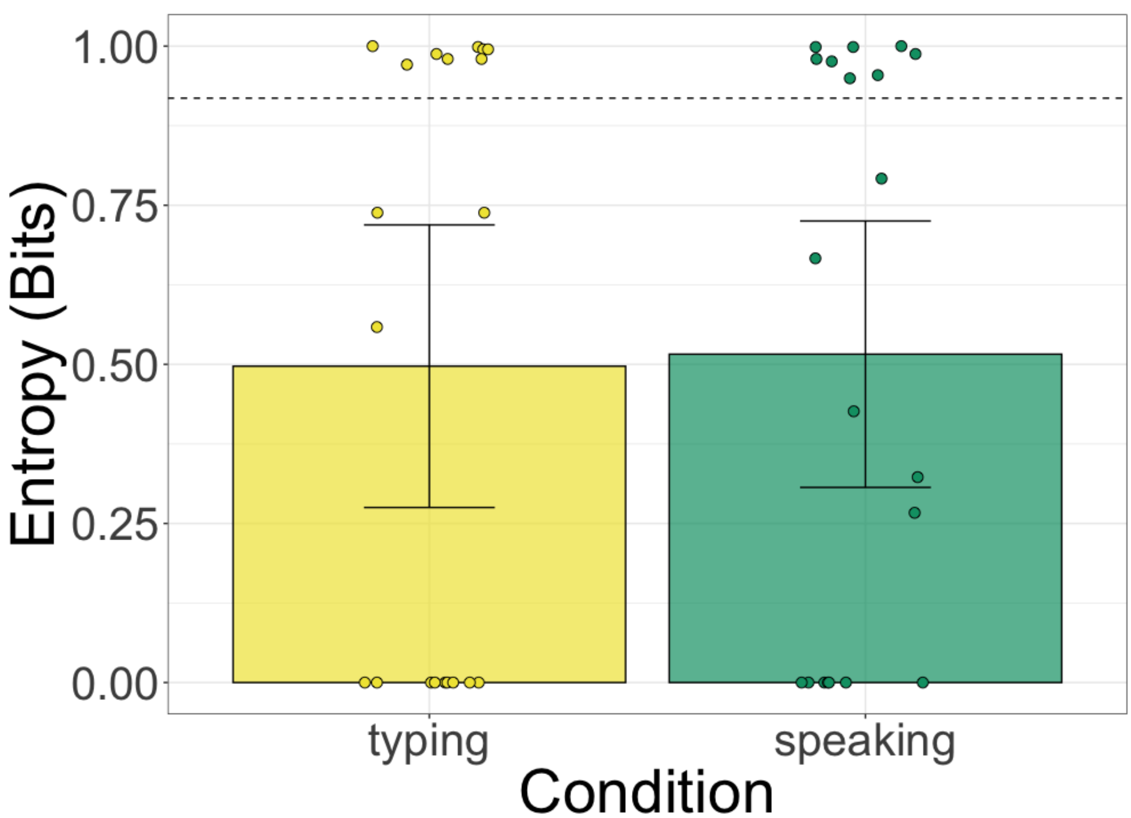

Figure 2.4: Entropy in participants' productions. The dashed line represents the 0.918 bits input language entropy. Dots represent individual participants' entropy. Error bars are bootstrapped 95% confidence intervals.

2.4 Discussion

We exposed participants to a miniature language with variable constituent order and measured whether their regularization behavior depended on the production modality of the experiment. We found no significant difference in the entropy of the output languages of participants who typed their productions into a textbox and participants who spoke their productions into a microphone. Thus, we did not find evidence that typing productions in a miniature language learning task leads to a difference in the regularization of variable constituent order.

While participants in the two production modality conditions did not differ from each other in amount of regularization, participants regularized the input language overall, for a reduction of entropy of 0.421 bits in the typing condition and 0.402 bits in the speaking condition. Does this mean that all participants had difficulty tracking variability in the miniature language? Perhaps this is the case. Other studies using a similar procedure took place over multiple days, giving participants extra exposure to the language and more time to master it (Fedzechkina et al. 2017, 2018). The current study's miniature language was less complex than these cited studies as we did not include linguistic elements such as case markers and prepositions, so we assumed that one session would be sufficient to learn the miniature language. It would be interesting to repeat this study over multiple days and see if participants regularized less as they became more proficient with the miniature language. There is likely another reason for the reduction of entropy observed in this study. Adults' regularization behavior appears to be dependent on their understanding of the task. Perfors (2016) found that adult participants matched variable miniature language input when they believed that the variation was predictable or when they believed that the goal of the miniature language learning task was to faithfully learn the variation, but they regularized the language when these task assumptions were removed. It is plausible that our participants did not interpret their task to be to faithfully reproduce variation and instead understood their goal to be communicative. If this is the case, participants could have eliminated unpredictable variation to facilitate communication (see: Smith & Wonnacott 2010). If speakers lower the entropy of a language, those with whom they are trying to communicate are less uncertain about the intended meaning of an utterance (Jaeger 2013, Kurumada & Jaeger 2015, Currie-Hall et al. 2018). For example, if a participant in our experiment produces an output language that has more VSO than the input language, the entropy is reduced, and any potential interlocutors will be more certain that the subject of the sentence is the first noun uttered. There is good reason to believe that participants in our study had a communicative goal in mind as participants in other studies with similar methodology restructured the input languages in ways that facilitate efficient communication and take into account social biases, indicating that this experimental paradigm encourages participants to treat the task as communicative (Fedzechkina et al. 2012, 2017, 2018, 2022, Roberts & Fedzechkina 2018).

An unexpected difference we saw between production modalities was the lexical error rate. Participants who spoke their productions made more lexical errors than participants who typed their productions in noun comprehension and noun production. The difference in amount of errors is also evident in the exclusion rate, as twelve participants in the speaking condition versus three in the typing condition were excluded from analysis for having more than 30% uncodeable trials at sentence production. These differences may be due to the additional written exposure to the miniature language that participants in the typing condition received. By introducing written forms paired with auditory input, participants in the typing condition received multimodal input when learning the language while participants in the speaking condition received auditory input alone. It would be interesting to repeat this study with the speaking condition receiving the same written input and seeing if the lexical error rate improves.

In our study, no significant differences in regularization behavior were found between participants who typed and spoke their productions, but participants who typed their productions made fewer lexical errors and fewer typers were excluded from analysis for having uncodeable utterances. These results, combined with the practical considerations of running an experiment over the web, lead us to recommend a typing production modality over a speaking production modality in miniature language learning experiments with adult participants.

Chapter 3

A Crosslinguistic Bias to Balance Production Effort Against Message Uncertainty[†]

3.1 Introduction

Over the past decade, researchers have used the results of miniature language learning studies as evidence that the structures of natural languages are shaped by general cognitive biases as participants in these studies prefer miniature language patterns that are crosslinguistically prevalent in natural languages (Culbertson et al. 2012, Fedzechkina et al. 2012, 2017, 2018, Kanwal et al. 2017, Kirby et al. 2015, Kurumada & Grimm 2019). For example, Fedzechkina and colleagues (2017) exposed participants to miniature languages comprised of sentences that had optional case marking (67% of sentences had a marker that denoted which element was the object while 33% did not) and either fixed (all sentences were in subject-object-verb order) or flexible (50% of sentences were in subject-object-verb order and 50% of sentences were in object-subject-verb order) constituent order. When participants used the miniature languages, those exposed to the flexible constituent order preferred using the case marker while those exposed to the fixed constituent order preferred not using it. This mirrors a crosslinguistically-prevalent pattern: languages with fixed constituent order tend not to have case marking while languages with flexible constituent order do (Blake 2001, Futrell et al. 2015, Sapir 1921)[1]. The researchers explain this effect in terms of a general cognitive bias thought to underlie language use: a bias to balance production effort against message uncertainty. The premise

[†]Results from English and German speakers were previously published in Hall Hartley & Fedzechkina (2020).

[1]In a quantification of constituent order flexibility, Futrell et al. (2015) found that nine out of 18 languages with low constituent order flexibility (subject-object relation order entropy below 0.50) had no case marking on nouns while only two out of 16 languages with high constituent order flexibility (subject-object relation order entropy above 0.50) had no case marking on nouns

of this bias is that speakers balance two opposing pressures on conveying their message: the pressure to expend as little effort as possible to convey the message (e.g., uttering the shortest sentence possible) and the pressure to make the message as informative as possible so that the listener understands (e.g., adding more elements to the sentence) (Jaeger 2010, Hawkins 2004, Piantadosi et al. 2011b, Zipf 1949)[2]. In Fedzechkina et al.'s study, the subject in the fixed order language always comes before the object, while in the flexible order language, there is more uncertainty about who is doing what to whom. The case marker takes effort to produce in both languages but does not provide much more information about grammatical function assignment in the fixed order language. Thus, participants who learn the fixed order language are less likely to use the case marker. This result has been replicated in Fedzechkina & Jaeger (2020) and in a related study, adults learning fixed constituent order miniature languages learned equally well whether or not a case marker was present, further supporting the idea that a case marker in a fixed constituent order language is redundant (Tal & Arnon 2022).

However, participants in these studies were adults who all speak a first language (L1). In studies of second (L2) and third (L3) language acquisition, there is often crosslinguistic influence from the L1, which may similarly be at play in these miniature language learning studies. L1 influence can be direct, such as English learners whose L1 has the same adverb-verb ordering as English noticing adverb-verb ordering errors in English more frequently than those whose L1 has the opposite ordering (Westergaard et al. 2017), or it can be more abstract, such as the presence of a vowel length contrast in the L1 facilitating a consonant length contrast in the L2 (Pajak & Levy 2014) or L2 learners of a language with case marking whose L1 uses a different case marking system making fewer case errors than learners whose L1 lacks a

[2] This bias fits into the Ideal Speaker Framework laid out by Kurumada & Jaeger (2015). In this framework, comprehension by the listener is a process of making an inference out of a noisy input, and the speaker can produce linguistic signals that differ in the degree to which the listener is likely to recognize the intended meaning. The ideal speaker is sensitive to feedback from the listener as well as the ease of production, and balances these two pressures in selecting their utterances.

case system altogether (Hopp 2010). Although investigators take precautions to not introduce direct L1 influence in miniature language learning studies (Fedzechkina et al. 2016), it is possible that an abstract influence from the L1 is responsible for the miniature language learning results mentioned above. Specifically, in Fedzechkina et al. (2017), participants were speakers of English, a language with relatively fixed constituent order and no morphological case marking on nouns. Although the constituent orders in the miniature languages were not English-like, English-speaking participants could have been dropping case in the fixed order miniature language because of a bias to use the miniature language in the way that they use their L1 (i.e., with fixed constituent order and no case), rather than a bias to balance production effort against message uncertainty. To test the possibility that using less case in a fixed order language is due to abstract influence from the L1, we expose participants whose L1s of German and Russian have case marking and flexible constituent ordering to the same miniature languages as English speakers and compare the three groups' miniature language use.

3.1.1 Participant L1s

English, German, and Russian were chosen as the L1s of interest for this study because the three languages use constituent order and case marking in different ways. English does not use case markers on nouns to denote the subject and object of the sentence (Leech 1975). Instead, English is a language with fixed subject verb object (SVO) constituent order, allowing speakers to determine grammatical function assignment from the order of nouns in the sentence (i.e., the subject usually precedes the object). Consider the sentences:

(1) a. The lion bites the man.

b. The man bites the lion.

In both sentences, an English speaker knows that the subject typically comes before the object and can use this fact to decide that the lion is doing the biting in (1a) and the man is doing the biting in (1b). German has a greater degree of constituent order flexibility than English, but obligatory case marking adds information about grammatical function assignment (Lederer 1969). In German, a sentence like *The lion bites the man* is typically in SVO order (2a), but object-verb-subject (2b) order is possible for emphasis:

(2) a. Der Löwe beißt den Mann.
The.NOM lion bites the.ACC man.
'The lion bites the man.'

b. Den Mann beißt der Löwe.
The.ACC man bites the.NOM lion.
'The lion bites the man.'

German speakers can infer that the lion is the one doing the biting in both sentences because the word for lion, *Löwe*, is preceded by a determiner with nominative case (marking the subject), *der*, while *Mann* is preceded by *den*, a determiner with accusative case (marking the object).

Of the three languages compared here, Russian has the greatest degree of constituent order flexibility (Futrell et al. 2015). Russian also has the richest case system of the three L1s, with six morphological cases used for nouns compared to German's four (der Dudenredaktion 1998, Wade 1992). Depending on what part of the sentence is being emphasized, the sentence *The lion bites the man* in Russian can have the words for lion and man appear on either side of the verb as in German—SVO (3a) and OVS (3b)–but verb-initial and verb-final orders are also possible, for example subject-object-verb (3c):

(3) a. Лев кусает мужчин-у.
Lion bites man-ACC.
'The lion bites the man.'

b. Мужчин-у кусает лев.
Man-ACC bites lion.
'The lion bites the man.'

c. Лев мужчин-у кусает.
Lion man-ACC bites.
'The lion bites the man.'

No matter the order of words in the sentence, Russian speakers can infer that the lion is the one doing the biting and the man is the one getting bit because the final *у* in *мужчину* (man) marks the object.

3.1.2 Predictions

Other miniature language learning studies using this method of crosslinguistic comparison have yielded mixed results. For example, Martin & Culbertson (2020) compared the similarity judgments of English (a suffixing language where inflectional morphemes are typically added to word endings) and Kîîtharaka (a prefixing language where inflectional morphemes are typically added to word beginnings) on a well-known miniature language learning task. Participants were presented with a sequence of geometric shapes or syllables and were asked to decide if a sequence that had extra elements added to the beginning (prefixed sequence) or a sequence with extra elements added to the ending (suffixed sequence) was more similar to the original sequence. In a previous study that tested English speakers only, Hupp et al. (2009) found that participants judged suffixed sequences as more similar to the first sequence and used this result to posit that a general cognitive preference for identifying sequences based on their beginnings was responsible for the crosslinguistic pattern that many times more languages are predominantly suffixing than are predominantly prefixing[3]. However, Martin & Culbertson found that this result did not hold for Kîîtharaka speakers, who judged prefixed sequences as more similar to the original sequence. Thus, performance in this task is influenced by the L1 of participants. On the other hand, another study using crosslinguistic comparisons did not find evidence that miniature language learning preferences were L1-dependent. Culbertson et al. (2020) recruited speakers whose L1s (French and Hebrew) order adjectives after nouns and numerals before nouns (non-harmonic ordering) and tested them on a miniature language learning task. In previous studies, Culbertson and colleagues found that speakers of English, a language with both adjectives and

[3]In a survey of 969 languages, 406 predominantly use suffixes and 58 predominantly use prefixes in their inflectional morphology (Dryer 2013)

numerals appearing before nouns (harmonic ordering), preferred to use harmonic orders to describe objects (2012, 2015). This is consistent with the crosslinguistic pattern that more languages use harmonic ordering than non-harmonic ordering. In this study with French and Hebrew speakers, Culbertson et al. (2020) found that even though their L1s used non-harmonic ordering, these speakers also preferred to use harmonic ordering in the miniature language. In this study, the results show a general preference for harmonic ordering, regardless of whether harmonic ordering is used in the L1 and gives further evidence for the crosslinguistic preference for harmonic ordering being due to a general cognitive bias.

Because of the mixed outcomes of prior investigations using this crosslinguistic comparison methodology, it is important to question if other well-known miniature language learning results are actually due to L1 influence. However, the mixed results discussed above also mean that it is unclear to us what to expect in our study. We lay out several possibilities here. If L1 influence is not responsible for the results of Fedzechkina et al. (2017), we expect all three L1 groups to show the same preference as English speakers in prior work: using more case marking in the flexible order miniature language where it adds information to the sentence than in the flexible order language where it is redundant. If an L1 influence is responsible for previous results, however, we expect to see some difference between L1 groups based on the fact that the three L1s use constituent order flexibility and case marking differently. For example, if our results are entirely dependent on L1, we might expect speakers of languages with case marking to retain case marking in the fixed order language. This would bring the fixed constituent order miniature language closer to both German and Russian where case marking on nouns is obligatory. Alternatively, if the results are a mix of L1 influence and a general bias to balance production effort against message uncertainty, we may see German and Russian speakers following the bias in a different way than English speakers. In the Fedzechkina et al. studies, participants did not change the word order input distribution of 50%/50% in the flexible order language (2017, 2020). It is possible that German and Russian speakers, who are familiar with word order variation, may be more comfortable changing this distribution. Increasing the use of one word order brings the miniature language closer to the fixed order language, where grammatical function assignment can be determined by the ordering of elements in a sentence and case marking is a redundant cue. If the German and Russian speakers drop case use in both the fixed order miniature language and the flexible order miniature language that they have made more fixed, they would be following the bias to balance message uncertainty against production effort (by reducing production effort when message uncertainty is low) in addition to following an L1-induced bias to change a constituent order distribution. In this study, we aim to further the understanding of results in miniature language learning studies. Our investigation will allow us to see whether the use of case marking and constituent order flexibility in miniature languages depends on how those devices are used in participants' L1s or if a general bias to balance message uncertainty against production effort operates without L1 influence.

3.2 Methods

3.2.1 Participants

The study consisted of two sessions administered on consecutive days via FindingFive, an online platform for behavioral study administration (FindingFive Team 2019). 168 participants were recruited from the crowdsourcing platform Prolific. Based on prescreening information from Prolific, all participants were native speakers of either English, German, or Russian (56 English, 61 German, and 51 Russian L1 speakers) with no language related disorders and an approval rate of 95% or higher based on at least 10 previous submissions on the platform. Participants received \$6.50 for completing the first session of the study which took an average of 47 minutes and an additional \$6.00 for completing the second session which took an average of 38 minutes. Based on prior work (Fedzechkina et al. 2017, Fedzechkina & Jaeger 2020), participants were recruited until there were 20 successful learners of each miniature language (fixed or flexible constituent order, see Section 3.2.2) from each L1 background (English, German, or Russian) for a total of 120 participants in analysis[4]. Successful learning was defined as scoring above 70% accuracy for unambiguous (i.e., case-marked) trials in the final sentence comprehension test on Day 2 of the study (as in Fedzechkina & Jaeger (2020)), and having more than 50% codeable sentence production responses on each day of the study (see Section 3.3.1). 21 participants (four English speakers and three German speakers learning the fixed order language; four English speakers, seven German speakers, and three Russian speakers learning the flexible order language) had at least 50% uncodeable responses on the first day of training and were not invited back to the second session. 12 participants were invited back to the second session but did not participate (four German speakers and one Russian speaker learning the fixed order language; five German speakers and two Russian speakers learning the flexible order language). 13 participants (eight English speakers, one German speaker, and four Russian speakers, all learners of the flexible order language) scored below 70% accuracy on the final sentence comprehension test on Day 2 and were thus excluded from analysis. Additionally, two participants (one German speaker and one Russian speaker, both learning the flexible order language) were excluded from analysis for providing more than 50% uncodeable responses on the second day of the study.

3.2.2 Input Miniature Languages

Participants learned either a fixed or flexible constituent order miniature language. Sentences in the fixed constituent order language were always in verb-subject-object (VSO) order. 50% of sentences in the flexible constituent order language were in VSO order and the other 50% were in verb-object-subject (VOS) order. These constituent orders were chosen because they highly uncommon in all of the participants' L1s[5]. In both input languages, the suffix *-dak* marked the object in 67% of the sen-

[4]post-hoc power analysis revealed that this sample size was sufficient with all effects having power > 0.84

[5]In English, all simple transitive sentences of the type used in the experiment are in subject-verb-object (SVO) order (Leech 1975). In German, most simple transitive sentences are in SVO

tences. The remaining 33% of sentences were not case marked. The case marker (when present) provided more information about grammatical function assignment in the flexible order language, where constituent orders were in free variation, than in the fixed order language where only one constituent order was used. Each noun occurred equally frequently as the object and subject of sentences. Each verb was used equally often with each constituent order possible in the miniature languages. Thus, grammatical function assignment could not be determined by the properties of the nouns or verbs alone. The two miniature languages used in the experiment contained the same lexicon of six nouns referring to animate alien characters ("bandit", "chef", "conductor", "hunter", "mountie", "referee"), two verbs referring to the transitive actions ("hug" and "kick"), and one case marker, optionally present as a suffix on the object noun of sentences, represented as *-dak*. All lexical items were phonotactically legal non-words of English, German, and Russian (Table 3.1). To ensure there were no unintended associations between lexical items and constituent order or case marking, the lexical items were rotated across two lists so that half of the participants were exposed to List A lexical item-referent pairs and half were exposed to List B. Nouns[6] and verbs were synthesized individually using the Greek-accented Melina voice of the Mac speech synthesizer. This voice was chosen to keep the accent in the miniature languages unlike any of the L1s of our participants. The individual words were concatenated into three-word sentences with 35 ms of silence between each word using Praat. This was done to ensure no prosodic cues to grammatical function assignment were available in our miniature languages. Participants were presented with words and sentences in the miniature languages both auditorily and in written form (necessary to familiarize participants with the novel orthography for typing their production responses, see Section 3.2.3). Written forms were

order, although object-verb-subject (OVS) is possible (Lederer 1969). In Russian, constituent order is more flexible, but VSO and VOS orders are rare (Bivon 1971).

[6]Both non-case-marked nouns (e.g. *balti*) and case-marked nouns (e.g. *balti-dak*) were synthesized individually before being concatenated with other nouns and verbs into sentences. This way, rather than being a separate element occurring 35 ms after the noun, the case marker sounded like a suffix on the noun.

Table 3.1: Miniature language lexicon in Latin script (Cyrillic script in brackets)

Nouns	balti [болти], dokla [докла], flaki [флаки], kusa [куса], mitsa [мица], nefli [нефли]
Verbs	dopsan [допсан], sunid [сунид]
Case marker	-dak [-дак]

presented in participants' L1 writing system (Latin script for English and German participants, Cyrillic script for Russian participants).

3.2.3 Procedure

The procedure was based on Fedzechkina & Jaeger (2020). Participants were instructed that they would be learning an alien language from an alien informant named Lumi by watching videos and hearing/reading their descriptions. They were told that near the end of the experiment, they would have a chance to write their own sentences in the alien language. A demonstration of the experiment can be watched at `https://youtu.be/2VH1xDtBhH0`. All instructions were provided in the participants' L1. Participants completed two sessions with at least 12 hours in between. The procedure was identical for both sessions. On each day, participants completed 12 blocks of training and testing, described below (Figure 3.1).

Noun Exposure

Participants first learned the names of the characters in isolation. They saw pictures of each referent accompanied by the corresponding noun in the miniature language in both auditory and written form (12 trials, two for each noun). The noun exposure block was followed by noun comprehension (Figure 3.2).

Noun Comprehension

Participants were presented with a noun in the miniature language and were instructed to select the character it referred to from a choice of four pictures of alien characters. Participants received feedback as to whether or not they had selected the correct image (12 trials, two for each noun). The noun comprehension block was followed by noun production.

Noun Production

Here, participants were presented with an image of the alien character and had to type the corresponding noun into a textbox. After they submitted their answer, participants saw and heard the correct noun associated with the image (six trials, one for each noun).

Participants completed two sequences of a noun exposure, noun comprehension, and noun production block before moving on to sentence exposure.

Sentence Exposure

Participants saw videos of two alien characters engaging in simple transitive actions (Figure 3.3). The videos were paired with sentences (presented auditorily and in written form) in the miniature language that described the scene. Participants could replay the video and sound as many times as they liked (24 trials total). Participants completed two sentence exposure blocks and then moved on to the sentence comprehension test.

Sentence Comprehension

Participants were presented with a sentence in the miniature language and had to select which of two videos it described. The videos differed in which alien character was the subject and which was the object. Participants could replay the audio and videos as many times as they wanted. No feedback was provided in sentence comprehension (24 trials total). After the first sentence comprehension block, participants engaged in two more sentence exposure blocks and a second sentence comprehension block before the final sentence production block.

Sentence Production

Participants were presented with previously unseen videos of two alien characters engaging in simple transitive actions. Participants were given verb prompts (verb in isolation presented auditorily and in written form in order to facilitate production) and were asked to write sentences in the miniature language that described what was happening in each video[7]. Participants could replay the audio and video as many times as they wanted. No feedback was provided (24 trials total).

Language Survey

At the end of the experiment, participants were provided with a textbox to answer the following question: "Please enter all languages you speak, the age (in years) at which you began speaking each, and your current proficiency (select between Low, Fair, Good, Excellent, Native) for each. *Example: English, from birth, Native; German, 16, Fair*"

[7]In piloting this study, we found that our initial instructions for the sentence production block were unclear. Many participants would only write the character names, which meant the utterance could not be analyzed for constituent order, and some participants would describe the video in their L1. We expanded the instructions to be more explicit, which led to fewer exclusions in the final study. We provide the instructions we used for sentence production in the final study in Appendix C.1.1.

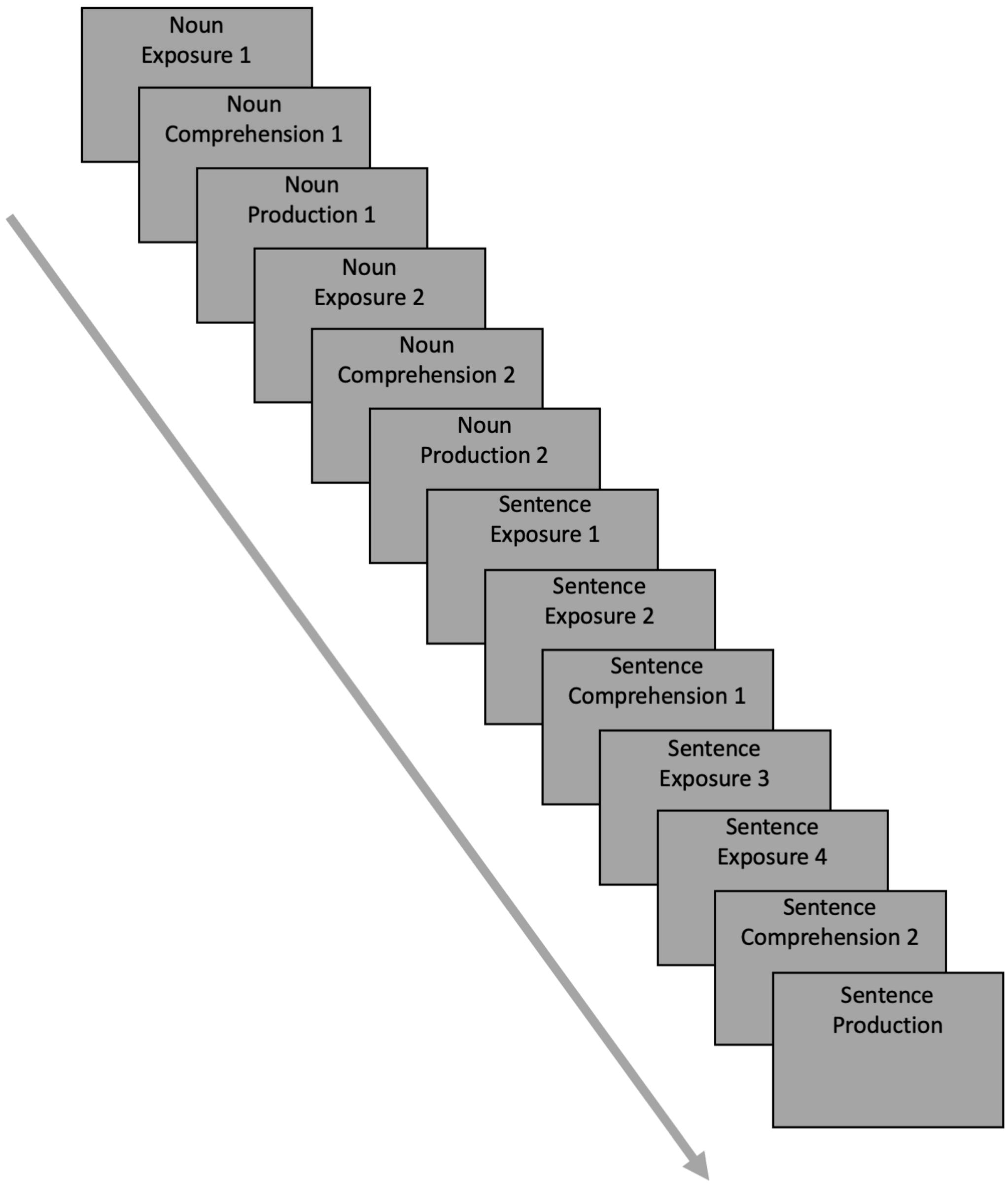

Figure 3.1: Overview of procedure blocks

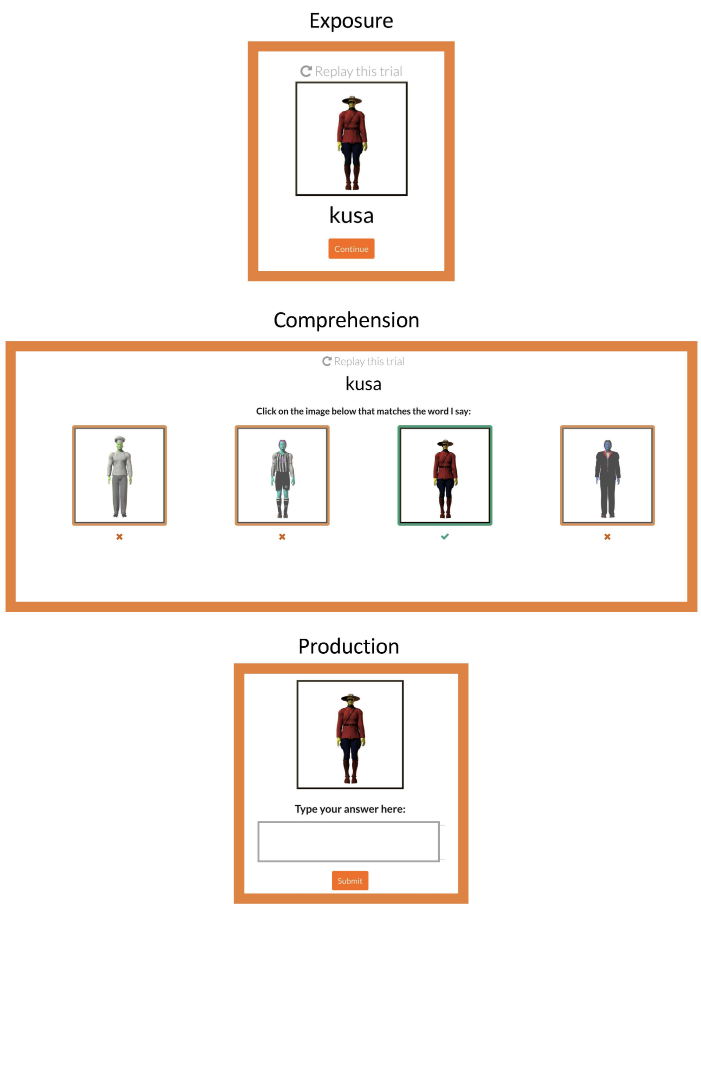

Figure 3.2: Screenshots from noun trials in English L1 condition.

Figure 3.3: Screenshots from sentence trials in English L1 condition.

3.3 Results

We are interested in whether miniature language learning performance is subject to L1 influence. Specifically, our main question of interest is whether the bias to balance message uncertainty against production effort is independent of L1. We investigate this question by comparing how participants of different L1 backgrounds use the devices of constituent ordering and case marking in their miniature language productions. In our setup, however, L1 influence on miniature language learning can manifest in a different way as well – as differences in learning accuracy between the L1 groups. Thus, before turning to the main question of whether learners express the bias to balance message uncertainty against production effort differently depending on their L1, we briefly describe how our data is scored, the results of the language survey conducted at the end of the study, and discuss learning accuracy across the three L1 groups.

3.3.1 Scoring

Sentence comprehension accuracy was calculated over case-marked trials, as these trials are unambiguous in both the fixed and flexible miniature language. Lexical accuracy in noun and sentence production trials was measured using soft-string matching with a custom Python script to calculate Levenshtein distance (the number of insertions, deletions, and substitutions) between a given word and the target word. The target words were the written lexical items presented in the experiment[8]. For all participants, any word where the Levenshtein distance was greater than two was coded as a lexical mistake. For example, the production *nifa* for target *nefli* was coded as a lexical mistake because it involves two substitutions and a deletion (Levenshtein distance = 3), while the production *nifli* was coded as a

[8] Russian speakers were presented with lexical items written in Cyrillic script, but a small number of Russian-L1 participants did not have access to a Cyrillic keyboard and typed their productions in Latin script. For these participants, target words were determined based on their consistent individual transliterations of the lexical items they saw in the Cyrillic script during the experiment.

correct response because it only involves one substitution (Levenshtein distance = 1). For sentence production trials, we also coded constituent order used in a sentence, presence of case marking, and where case was marked (i.e., on the subject, object, or verb). Grammatical mistakes were defined as using a constituent order not present in the input or using the case marker on an element other than the object. Productions that contained grammatical mistakes (0.9% of the data on the final day of training) were removed from analysis. Lexical mistakes (5.0% of the data on the final day of training) were kept in all of the analyses as long as sentence constituent order and the presence of case marking could be reliably determined (i.e., when a sentence contained only one lexical mistake). Sentence production responses containing more than one lexical mistake were marked as uncodeable and were excluded from analysis (0.7% of the data on the final day of training).

3.3.2 Language Survey Results

Most English L1 speakers included in analysis were monolingual. By virtue of having signed up for a Prolific account in English, we expected German and Russian L1 speakers to have some understanding of English. Most German and Russian L1 speakers included in analysis reported at least "Good" (see Section 3.2.3) proficiency in English. Several German and Russian L1 speakers also reported at least "Good" proficiency in other languages, including one German L1 speaker with proficiency in Russian and four Russian L1 speakers with proficiency in German (Table 3.2).

3.3.3 Learning Accuracy

Sentence Comprehension Accuracy

Here we ask whether there are differences in sentence comprehension accuracy between our three L1 groups. We calculated mean percent sentence comprehension accuracy across unambiguous (case-marked) trials (Table 3.3). Overall, sentence comprehension accuracy is quite high (97.8% average accuracy across groups on the final block of training). We might expect that participants with morphological case

Table 3.2: Number of participants reporting "Good" or "Excellent" proficiency for languages other than their L1

Other Language	English L1	German L1	Russian L1
Dutch		1	
English		39	36
Finnish			1
French	1	7	1
German			4
Greek			1
Hebrew			3
Italian			5
Kazakh			1
Kyrgyz			1
Latin		1	
Latvian			2
Luxembourgish		1	
Norwegian		1	
Polish			1
Russian		1	
Spanish		1	2
Swedish			1
Ukrainian			2

marking on nouns in their L1 (German and Russian speakers) learn the miniature language case system more accurately than English speakers, in which case their sentence comprehension accuracy would be higher. We do find that English speakers had the lowest sentence comprehension accuracy out of participants learning both the fixed and flexible miniature languages on the first block of the first day of training and on both blocks of the second day of training. However, confidence intervals overlap when comparing the three L1 groups for each miniature language. Thus, the effect of having a morphological case system on improved sentence comprehension accuracy in this task is small at best.

Table 3.3: Mean percent accuracy for unambiguous sentence comprehension trials (95% confidence intervals shown in brackets)

		English		German		Russian	
Day	Block	Fixed	Flexible	Fixed	Flexible	Fixed	Flexible
1	1	87.02 [75.74, 95.31]	77.81 [64.68, 89.38]	94.38 [87.19, 98.75]	85.31 [74.99, 94.38]	96.56 [91.55, 99.69]	79.38 [70, 88.13]
	2	95.94 [93.13, 98.13]	96.88 [92.19, 99.69]	98.44 [97.19, 99.38]	89.69 [77.19, 97.81]	95.94 [90.13, 99.38]	90.94 [83.75, 96.56]
2	1	99.38 [98.13, 100]	95 [89.69, 98.75]	99.38 [98.44, 100]	96.56 [92.5, 99.38]	98.75 [97.5, 99.69]	96.88 [93.74, 99.06]
	2	96.88 [93.75, 99.38]	96.56 [93.75, 98.75]	99.69 [99.06, 100]	98.75 [97.19, 100]	97.5 [95.94, 99.06]	97.5 [95.63, 99.06]

Grammatical Mistakes in Sentence Production

As German and Russian have noun case marking and constituent order flexibility, we might expect German and Russian speakers to make fewer case errors in production of both miniature languages and fewer constituent order errors in the flexible order miniature language. Overall, grammatical mistakes in sentence production were low across all L1 groups (1.07% average across both days and all groups). Case errors, defined as sentences with case marking on the subject or verb rather than the object, were numerically lower for German L1 speakers than for English and Russian L1 speakers on Day 1, but on Day 2, the English and Russian L1 speakers had fewer case errors than the German speakers (Table 3.4). Thus, we see no clear pattern between the presence of an L1 case system and number of case errors in miniature language production. Sentences with case errors made up only a small percentage of total sentences across both miniature languages and all L1s (1.1% on Day 1 and 0.4% on Day 2) so we were unable to do further reliable analysis. Constituent order errors, defined as sentences that used a constituent order not present in the input, were numerically highest for English speakers learning the fixed order language on Day 1 and highest for German speakers learning the fixed order language on Day 2 (Table 3.5). The high proportion of constituent order errors for English speakers in the fixed order language on Day 1 is somewhat surprising, as English is a fixed order language and we might expect English speakers to consistently use one constituent order (VSO) in their productions. Sentences with constituent order errors made up 2.1% of sentences on Day 1 and 0.5% of sentences on Day 2 (across both miniature languages and all L1 groups), which again is a small percentage of total trials.

The high accuracy on the sentence comprehension test and low rates of grammatical mistakes in sentence production suggest that our participants fully learned both miniature languages regardless of L1. Numerically, we found few learning differences between the L1 groups. Next, we look at our main factors of interest: constituent order and case use across the three L1 groups in miniature language production.

Table 3.4: Mean percent of sentences with case errors in sentence production (95% confidence intervals shown in brackets)

	English		German		Russian	
Day	Fixed	Flexible	Fixed	Flexible	Fixed	Flexible
1	1.14 [0.23, 2.28]	1.28 [0.43, 2.34]	0.21 [0, 0.63]	1.08 [0.22, 2.15]	1.27 [0.42, 2.54]	1.73 [0.65, 3.02]
2	0 [0, 0]	0.85 [0.21, 1.71]	0.42 [0, 1.04]	1.04 [0.21, 1.87]	0 [0, 0]	0.21 [0, 0.63]

Table 3.5: Mean percent of sentences with constituent order errors in sentence production (95% confidence intervals shown in brackets)

	English		German		Russian	
Day	Fixed	Flexible	Fixed	Flexible	Fixed	Flexible
1	7.06 [4.56, 9.57]	2.56 [1.28, 4.04]	1.47 [0.42, 2.52]	0 [0, 0]	1.91 [0.85, 3.18]	0 [0, 0]
2	0.21 [0, 0.63]	0 [0, 0]	2.71 [1.25, 4.17]	0.21 [0, 0.63]	0.42 [0, 1.05]	0 [0, 0]

3.3.4 Constituent Order Use in Sentence Production

The three L1s in our experiment allow different amounts of constituent order flexibility (English<German<Russian). Thus, we first ask whether participants display different preferences in constituent order use in miniature language productions depending on their L1. We look at the productions of participants who learned the flexible order miniature language only, as given our scoring, all participants who learned the fixed order miniature language trivially matched the input proportion of 100% VSO (see Section 3.3.1). We first compare proportion VSO use across the L1 groups. Then, we compare each L1 group's VSO use to the input. In Fedzechkina et al. (2017), English speakers matched the input proportion of constituent orders. German and Russian speakers are more familiar with constituent order flexibility, and thus may be more likely to change the VSO/VOS distribution from 50%/50% than English speakers.

To investigate whether participants of different L1s learning the flexible order language had different preferences in constituent order use, we fit a generalized linear mixed effects model that predicted VSO use from L1 (sliding contrast coded[9]; German compared to English, Russian compared to German), day of training (sum coded; 1=Day 2, -1=Day 1), and their interactions. The model contained the fullest converging random effects structure (by-participant random intercept). L1 groups did not differ in the amount of VSO used in production (German vs. English contrast: $\hat{\beta}$ = 0.064, z = 0.197, p = 0.843; Russian vs. German contrast: $\hat{\beta}$ = 0.468, z = 1.444, p = 0.148; full results in Table A.4). Participants across all L1 backgrounds used less VSO on Day 2 than on Day 1 ($\hat{\beta}$ = -0.128, z = -3.096, p = 0.001). There was an interaction between the German vs. English comparison and day of training ($\hat{\beta}$ = -0.288, z = -2.814, p = 0.004) but no significant interaction between the Russian vs. German comparison and day of training ($\hat{\beta}$ = 0.097, z = 0.965, p = 0.334). A simple effects test revealed that German ($\hat{\beta}$ = -0.257, z = -3.619, $p < 0.001$; full results in Table A.5) and Russian ($\hat{\beta}$ = -0.160, z = -2.246, p = 0.024) L1 speakers used significantly less VSO on Day 2 than on Day 1, while English L1 speakers did not ($\hat{\beta}$ = 0.030, z = 0.420, p = 0.674), suggesting that the effect of Day was driven by Russian and German speakers. This finding suggests that German and Russian speakers used less VSO as they gained more experience with the miniature language, while English speakers used about the same proportion of VSO constituent order on both days of the study.

We next asked whether English, German, and Russian L1 speakers learning the flexible order language deviated from the VSO constituent order input proportion (50%) on the final day of the experiment. We used a generalized linear mixed effects model that predicted VSO use from L1 (treatment coded) and the fullest converging random effects structure (by-participant random intercept). The intercept of this

[9]Sliding contrast coding compares the mean of the dependent variable of one level to the mean of the dependent variable of the level that comes before it. Here, we have ordered our levels as English<German<Russian as we expect any differences between the groups to depend on the richness of case systems or flexibility of constituent ordering, both of which English has the least and Russian has the most.

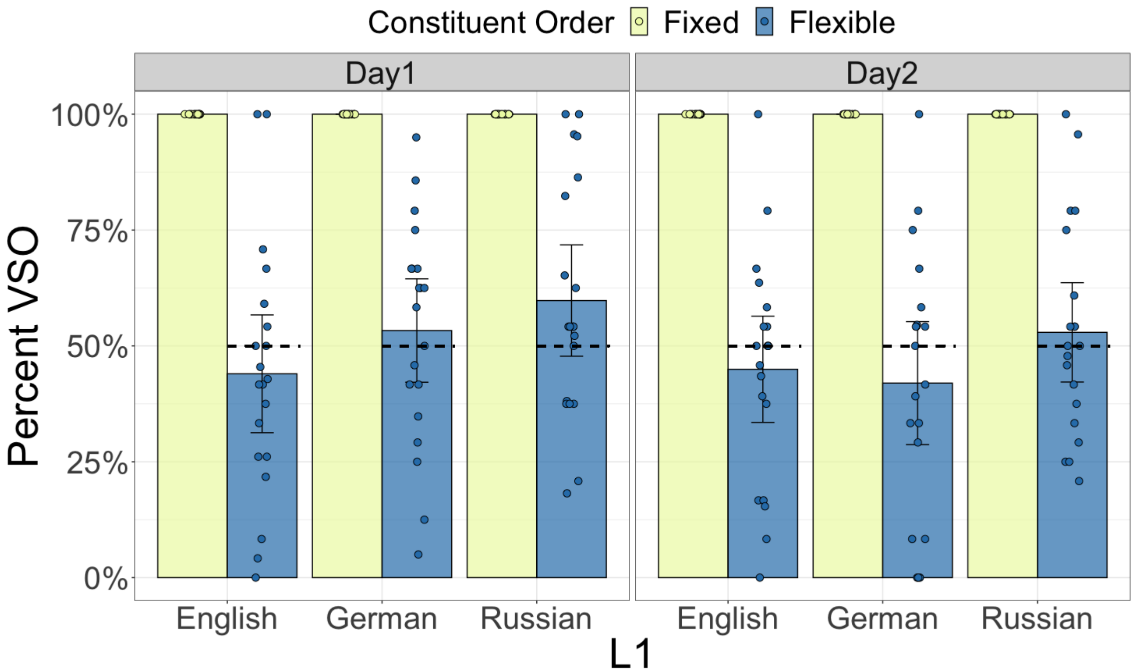

Figure 3.4: VSO use in sentence production by day of training, constituent order flexibility, and L1 background. The dashed line represents the 50% input percentage VSO for the flexible order language (VSO input for the fixed order language is 100%). Dots are individual participants' percentages of VSO. Error bars are bootstrapped 95% confidence intervals.

model captures whether the L1 group coded as the reference level differs from the 50% input proportion of VSO use. We ran this model three times, with each L1 coded as the reference level. We found that on the final day of the experiment, all L1 groups matched the input VSO proportion (English, 44.9%: $\hat{\beta}$ = -0.263, z = -0.826, p = 0.409, German, 41.5%: $\hat{\beta}$ = -0.510, z = -1.578, p = 0.115, Russian, 52.9%: $\hat{\beta}$ = 0.213, z = 0.674, p = 0.500; Figure 3.4). This finding replicates and extends Fedzechkina et al. (2017) showing that learners' preference to match the input proportions of constituent orders does not depend on L1.

Our two analyses of constituent order use show no overall difference between L1 groups in proportion VSO used in the flexible order miniature language. Even though German allows significantly more constituent order flexibility than English and Russian allows more than German, learners' L1s did not affect the proportion of VSO used in our study.

3.3.5 Case Use in Sentence Production

We now turn to our main research question. If the bias to balance message uncertainty against production effort does not depend on L1 structure, we expect to see all three L1 groups following this bias in the same way. Specifically, we expect learners of the fixed order language to use less case than learners of the flexible order language, as case marking takes effort to produce and only significantly reduces message uncertainty in the flexible order miniature language. If, however, the presence of a case system in the L1 influences case use in the miniature languages tested, Russian and German speakers, who have case systems in their L1, may use case differently than English speakers who do not have case in their L1.

To answer this question, we compared the proportion of case used by each L1 group in each miniature language by fitting a generalized linear mixed effects model that predicted case use from L1 (sliding contrast coded; German compared to English, Russian compared to German), miniature language constituent order flexibility (sum coded; 1=flexible order, -1=fixed order), day of training (sum coded; 1=Day 2, -1=Day 1), and their interactions. The model contained the fullest converging random effects structure (by-participant and by-item random intercepts, by-participant and by-item random slopes of day). Constituent order flexibility had a significant effect on case use: Across all L1s, learners of the flexible order language used significantly more case than learners of the fixed order language ($\hat{\beta}$ = 1.286, z = 6.344 $p < 0.001$; Figure 3.5; full results in Table A.6). German speakers did not differ in the overall amount of case from English speakers ($\hat{\beta}$ = -0.467, z = -0.949, p = 0.342), and Russian speakers used the same amount of case as German speakers ($\hat{\beta}$ = -0.729, z = -1.484, p = 0.137). There was a significant main effect of day, with participants using more case on Day 2 compared to Day 1 ($\hat{\beta}$ = 0.248, z = 2.952, p = 0.003). There was also an interaction between constituent order flexibility and day of training ($\hat{\beta}$ = 0.187, z = 2.334, p = 0.019). A simple effects test revealed that learners used significantly more case on Day 2 compared to Day 1 in the flexible

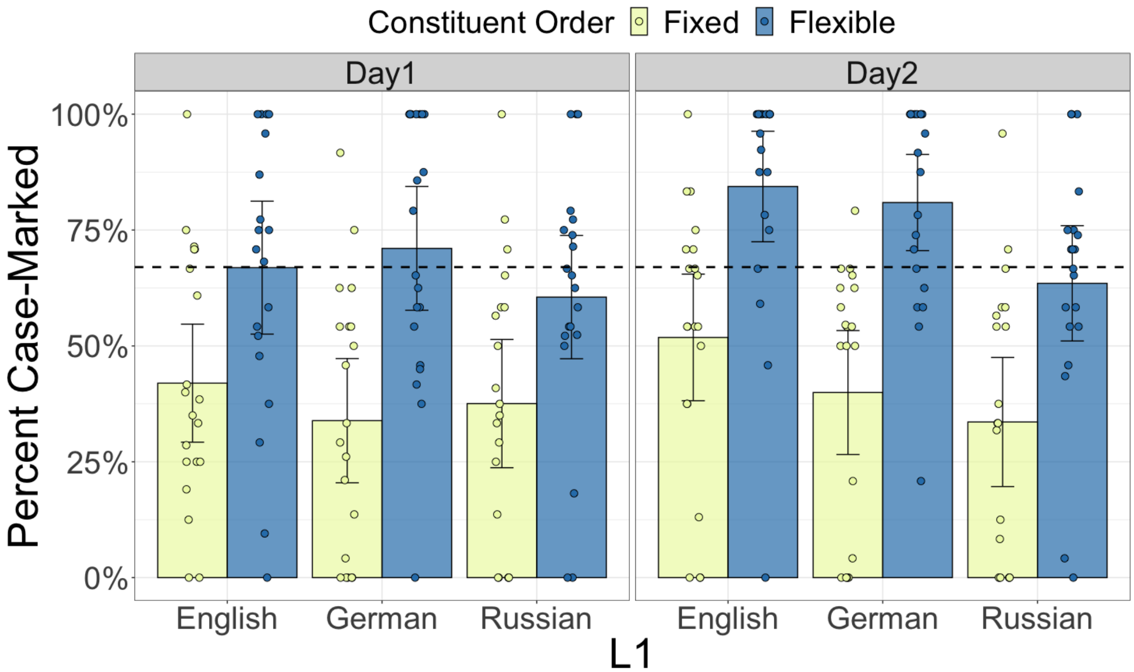

Figure 3.5: Case use in sentence production by day of training, miniature language constituent order flexibility, and L1 background. The dashed line represents the 67% input percentage of case (same for both miniature languages). Dots are individual participants' proportions case-marked productions. Error bars are bootstrapped 95% confidence intervals.

order language ($\hat{\beta} = 0.436$, z = 3.645, p < 0.001; full results in Table A.7) but not in the fixed order language ($\hat{\beta} = 0.060$, z = 0.539, p = 0.590). This increase in case use from Day 1 to Day 2 could suggest that as participants become more familiar with the language, they use more case to add information about grammatical function assignment as they realize that constituent order alone is not a reliable cue.

Interactions between L1 and day of training as well as three-way interactions between constituent order flexibility, L1, and day of training failed to reach significance (p's > 0.084). Notably, there was not an interaction between miniature language constituent order flexibility and L1 group for either the German vs. English contrast ($\hat{\beta} = 0.465$, z = 0.945, p = 0.344) or the Russian vs. German contrast ($\hat{\beta} = -0.657$, z = -1.337, p = 0.181), meaning that learners from all three L1 groups behaved in the same way using more case in the flexible order miniature language than the fixed.

On Day 2, we see that numerically, English speakers use the most case and Russian speakers use the least amount of case in both the fixed and flexible miniature language conditions. While the main effects of L1 did not reach significance in our model, the negative estimates for the German vs. English comparison and the Russian vs. German comparison show that Russian speakers use less case than German speakers, and German speakers use less case than English speakers. We put forth a possible explanation for this numerical trend in the discussion.

Overall, all three L1 groups use more case in the flexible order language, where case marking adds information about grammatical function assignment, than in the fixed order language, where case marking is a redundant cue. Thus, English, German, and Russian speakers follow the bias to balance production effort against message uncertainty in the same way.

3.4 Discussion

We add to the body of work investigating L1 influence on preferences found in miniature language learning studies with adults. Using miniature languages with optional case marking and either fixed or flexible constituent order, we compared the case marking presence and constituent ordering in miniature language sentences produced by participants who differ in how case and constituent order flexibility are employed in their L1. We find that English, German, and Russian speakers follow the bias to balance production effort against message uncertainty in the same way: They match the input constituent order proportions and use more case in the flexible order miniature language than in the fixed.

Differences between L1 groups in the learning phases of the experiment are small. English speakers numerically had the lowest sentence comprehension accuracy rates in the first block of the experiment, which we might expect if it takes speakers without case marking on nouns in their L1 longer to learn a case system in a miniature language. However, this difference does not hold as blocks progress, and 95% confidence intervals overlap for the accuracy scores between each L1 on each block.

Similarly, we did not see a clear pattern of any L1 group having more case errors and constituent order errors in the sentence production blocks of the experiment. As our study was designed to assess a language use bias which requires participants to have fully learned the miniature language, we saw very high accuracy rates that could mask learning differences that may exist between L1 groups. Thus, the question of if English, German, and Russian speakers learn miniature languages with fixed and flexible constituent ordering and optional case marking differently should be answered with a different paradigm that allows for the measurement of learning differences.

One small difference we did see between L1 groups was the amount of case used in sentence production. Numerically, English speakers used the case marker on more sentences than German speakers and German speakers used the case marker on more sentences than Russian speakers. Why would speakers of Russian, a language that has morphological case marking, use the least amount of case? The answer may lie in the number of similarities that the miniature language shares with the L1s. The miniature language where we see this effect has flexible constituent order, unlike English that has fixed constituent order, somewhat like German that has some flexibility in constituent ordering, and like Russian that has flexible constituent order. The optional case marker is a suffix on the object noun, which is also most like the Russian case expression and least like English's lack of case marking on nouns. Thus the decrease in case use follows as participants' L1s get more similar to the miniature language. Russian speakers also most closely match the input distribution of case marking. It is possible that the similarity between the flexible order miniature language and the L1 led Russian speakers to better understanding the miniature language than German speakers and German speakers to better understanding the miniature language than English speakers, leading Russian speakers to most closely follow the input distribution of the miniature language while English and German speakers were more driven by the bias to decrease message uncertainty in a language with variable constituent ordering. This effect brings up a question we attempt to answer in Chapter 4: If we increase the similarity between the miniature language and participants' L1, will participants abandon the bias for balancing production effort against message uncertainty?

It is important to note that 39/40 German L1 speakers and 36/40 Russian L1 speakers reported at least "Good" proficiency in English. This is a consequence of online recruitment from a platform where the majority of hosted studies are in the English language. However, we strove to limit any English influence for German and Russian L1 participants by presenting all recruitment materials and instructions in the participants' L1s in an effort to keep them in their L1 language mode for the study (Grosjean 2001, Marian & Spivey 2003, Canseco-Gonzalez et al. 2010). It would, however, be ideal to repeat this study with all monolingual speakers to ensure that the only crosslinguistic influence that could appear in the miniature language preferences of participants is from the L1.

The overall pattern of results in this study replicates and expands on previous work: participants use more case in a flexible constituent order miniature language than a fixed regardless of how constituent ordering and case marking are expressed in their L1. Our work supports the claim that the bias to balance production effort against message uncertainty does not depend on learners' L1, and that this bias may be responsible for the crosslinguistic trade-off of constituent order flexibility and case marking. By continuing to include speakers of structurally different languages in miniature language learning studies, researchers can tease apart behaviors that are due to L1 influence and those due to general cognitive biases.

Chapter 4

The Effect of Language Similarity on a Bias to Balance Production Effort Against Message Uncertainty

4.1 Introduction

In Chapter 3, we asked whether the preference for miniature language learners to drop case marking in a fixed constituent order language but retain it in a flexible constituent order language could be due to an abstract influence from the L1 rather than a general cognitive bias to balance production effort against message uncertainty. We did not find evidence that this preference was influenced by learners' L1s, as English, German, and Russian speakers whose L1s use case marking and constituent order flexibility in different ways all displayed the same miniature language use preferences. This result supports the idea that a bias to balance production effort against message uncertainty underlies the trade-off between constituent order flexibility and case marking seen crosslinguistically, with more languages with rigid constituent ordering lacking case marking than languages with flexible constituent ordering.

This result is somewhat surprising. Miniature language learning researchers typically try to make their miniature languages as different from learners' L1s as possible, often by investigating structures that are not present in the participants' L1 to avoid L1 influence affecting the results (Fedzechkina et al. 2016). In Chapter 3 however, our miniature languages marked case on the object noun as do German and Russian, but German and Russian L1 speakers did not show strong evidence of L1 influence. Why is this the case? Perhaps the behavior of using less case marking in a fixed order than flexible order miniature language is driven by a true general cognitive bias that is immune to L1 influence. Or, perhaps our miniature languages were still too different from participants' L1s to induce L1 influence. Detecting the similarity between the L1 and the target language for natural language acquisition appears to be an important factor in whether or not L1 influence arises (Rothman 2011).

Several theories of crosslinguistic influence in second and third language learning highlight the importance of *perceived similarity* between the languages a learner speaks and the target language (Pajak et al. 2016, Rothman 2015, Westergaard et al. 2017). That is, the factor determining if properties are transferred from the L1 when learning a new language is if the learner decides that there is similarity between the L1 and the target language (Pajak et al. 2016). The perceived similarity can be a top-down belief, for example the learner knowing that the target language is in the same language family as the L1 (Pajak et al. 2016), or can come from the input, such as detected similarity in the lexicon or in morphosyntactic structures (Rothman 2015). For example, take the case of Spanish-English bilinguals learning French studied by Rothman & Cabrelli-Amaro (2010). Spanish is a null-subject language while both English and French require subjects. Even though the bilinguals had knowledge of a non-null-subject language (English), they transferred the null-subject property of Spanish to the non-null-subject target language (French), presumably because the perceived similarity between the two Romance languages (Rothman & Cabrelli-Amaro 2010). Or consider Spanish-English bilinguals learning Brazilian Portuguese: A generic interpretation of bare plurals is possible in both English (*Cats are beautiful*) and Brazilian Portuguese (*Gatos são belos*) but not in Spanish (#*Gatos son bellos*). When asked to judge the acceptability of a generic interpretation of bare plurals in Brazilian Portuguese, Spanish-English bilinguals rated these sentences as unacceptable, again showing an influence from the other Romance language that the learners spoke (Iverson 2010).

In our study, it is possible that German and Russian speakers did not perceive the similarity between the miniature languages and their L1s because while we did include structures present in German and Russian, we still tried to make our miniature languages as unlike English, German and Russian as possible. We used verb-initial constituent orders, which are not used in English and German and are very rare in Russian (Leech 1975, Lederer 1969, Bivon 1971). Additionally, our case marker, *-dak* did not share formal similarity to German case markers, which are expressed on determiners preceding nouns unlike our suffixal marker, or to Russian case markers, which are conditioned on the phonology of the noun while ours was not. Thus, while the abstract similarity between German/Russian and our miniature languages was not enough to induce L1 influence, it is possible that more direct similarity would be.

Here, we test the hypothesis that L1 influence on miniature language learning depends on the degree of miniature language-to-L1 similarity. We use the same miniature language setup as Chapter 3 as a test case and ask whether learners continue to use less case in a fixed constituent order language than a flexible constituent order language as the languages get closer to learners' L1. We compare the miniature language learning preferences of German L1 speakers on three sets of fixed/flexible constituent order miniature languages with optional case marking that get increasingly similar to German: The languages from Chapter 3 with VSO/VOS constituent order and a suffixal case marker on the object (constituent ordering and case marker form unlike German), languages with VSO/VOS constituent order and a case marker as a determiner preceding the object (constituent ordering unlike German; case marker form like German), and languages with SVO/OVS constituent order and a case marker as a determiner preceding the object (constituent ordering and case marker form like German).

The results of this study will help elucidate why, to date, certain miniature language learning studies testing for L1 influence have not found an effect (Chapter 3, Culbertson et al. 2020) while others have (Martin & Culbertson 2020). If the degree of miniature language-to-L1 similarity influences learning preferences, researchers will have more knowledge about how to design miniature languages to test for general cognitive biases (i.e., make the miniature language as dissimilar to the L1 as possible). If no similarity effects are found, there will be stronger evidence for the possibility that only certain biases are susceptible to L1 influence and that the bias we investigate, that to balance production effort against message uncertainty, is resistant to such effects.

4.2 Methods

Here we compare the preferences of German speakers on the miniature languages in Chapter 3 to new miniature languages. Aside from the structure of the miniature languages, the methods are identical to Chapter 3.

4.2.1 Miniature Languages

We compare miniature language preferences on three sets of miniature languages that share increasingly more similarities to German. The **Verb-Initial Suffix** (VI-SFX) set of languages are those used in Chapter 3: The fixed constituent order language has 100% VSO order and the flexible constituent order language has 50% VSO and 50% VOS order. Both VI-SFX languages have 67% of input sentences marked with a suffixal case marker -*dak* attached to the object noun (Fig. 4.1). The **Verb-Initial Determiner** (VI-DET) set of languages are designed to be more similar to German by making the case marker *dak* a determiner that appears before the object noun. The fixed constituent order VI-DET language has 100% VSO and the flexible order language has 50% VSO and 50% VOS. Both VI-DET languages have 67% of input sentences marked with the determiner *dak* preceding the object noun. The **Verb-Medial Determiner** (VM-DET) set of languages are designed to be the most similar to German by again marking case on a determiner preceding the object noun but also by using verb-medial SVO and OVS constituent orders as in German simple transitive sentences. The fixed order VM-DET language has 100% SVO order and the flexible order language has 50% SVO and 50% OVS order. Both VM-DET languages have 67% of input sentences marked with the determiner *dak* preceding the object noun.

Similarity to German →

	VI-SFX	VI-DET	VM-DET
Fixed Order	100 % VSO	100% VSO	100% SVO
Flexible Order	50% VSO 50% VOS	50% VSO 50% VOS	50% SVO 50% OVS
German-like case marking		✓	✓
German-like constituent ordering			✓
Example	dopsan mitsa kusa-dak *hug chef mountie-CM*	dopsan mitsa dak kusa *hug chef CM mountie*	mitsa dopsan dak kusa *chef hug CM mountie*

Figure 4.1: Miniature languages used in Chapter 4. VI-SFX languages have the case marker as a suffix on the object noun while VI-DET and VM-DET languages have the case marker as a determiner preceding the object noun. The 'Example' row shows sentences describing the video screenshot on the left in the subject-before-object order with case marked for each set of languages.

4.2.2 Participants

We recruited participants to learn the VI-DET and VM-DET languages and compared the performance of these participants to that of the German L1 participants in Chapter 3. Following Chapter 3, we recruited self-reported native German speakers from Prolific who had no language related disorders and an approval rate of 95% or higher on at least 10 previous submissions on the platform. In total, 103 participants were recruited. Recruitment continued until we had 20 successful learners of each language: 20 in the fixed VI-DET language, 20 in the flexible VI-DET language, 20 in the fixed VM-DET language, and 20 in the flexible VM-DET language. As in Chapter 3, successful learning was defined as scoring above 70% accuracy for unambiguous trials in the final sentence comprehension test on Day 2 of the study and having more than 50% codeable sentence production responses on each day of the study. Eight participants (three leaning the VI-DET flexible language, three learning the VM-DET flexible language, and two learning the VM-DET fixed language) had more than 50% uncodeable responses on the first day of training and were not invited back to the second session. Three participants (one learning the VI-DET flexible language and two learning the VI-DET fixed language) were invited back to the second session but did not participate. Twelve participants (two learning the VI-DET flexible language, nine learning the VM-DET flexible language, and one learning the VM-DET fixed language) scored below 70% accuracy on the final sentence comprehension test on Day 2 and were thus excluded from analysis. 120 participants were included in analysis, 80 of whom were recruited for this study in 2021 (20 learned VI-DET fixed, 20 VI-DET flexible, 20 VM-DET fixed, and 20 VM-DET flexible) and 40 of whom were recruited for the study in Chapter 3 in 2020 (20 learned VI-SFX fixed and 20 learned VI-SFX flexible). Pay for the new participants was the same as in Chapter 3: $6.50 for the first session and $6.00 for the second.

4.3 Results

We compare the miniature language preferences from German L1 speakers who learned languages with VSO/VOS constituent order and a suffixal case marker on the object noun (VI-SFX languages; dissimilar to German constituent order and case marking), VSO/VOS constituent order and a determiner case marker preceding the object noun (VI-DET languages; dissimilar to German constituent order and similar to German case marking), and SVO/OVS constituent order and a determiner case marker preceding the object noun (VM-DET languages; similar to German constituent order and case marking). In each set of languages, 20 participants learned a fixed constituent order language where the subject always preceded the object and 20 participants learned a flexible constituent order language where the subject preceded the object half the time. We are interested to see if participants in all three sets of language match the input constituent order distribution and use less case in the fixed order language than the flexible as they did in the VI-SFX languages (suggesting a bias to balance production effort against message uncertainty that is robust against L1 influence), or if increasing the similarity to German causes participants to prefer to use the miniature languages differently.

Before our analysis of constituent order and case use, we first report language survey results, then look at learning accuracy for any differences between miniature languages learned.

4.3.1 Language Survey

As in Chapter 3, the German L1 speakers we recruited are mostly proficient in English (39/40 learning VI-SFX languages, 40/40 learning VI-DET languages, 39/40 learning VM-DET languages). Participants also reported "Good" or "Excellent" proficiency in Croatian, Dutch, French, Italian, Ladin, Latin, Luxembourgish, Norwegian, Polish, Portuguese, Russian, Spanish, and Swiss German (Table 4.1).

Table 4.1: Number of participants reporting "Good", "Excellent" proficiency for languages other than their L1

	VI-SFX	VI-DET	VM-DET
Croatian			1
Dutch	1		1
English	39	40	39
French	7	3	5
Italian		1	1
Ladin			1
Latin	1		1
Luxembourgish	1		1
Norwegian	1		1
Polish			2
Portuguese		1	
Russian	1	1	1
Spanish	1	4	1
Swiss German			1

4.3.2 Learning Accuracy

Sentence Comprehension Accuracy

We first ask if sentence comprehension accuracy depends on the miniature language participants are exposed to. We might expect that if the miniature language is more similar to participants' L1s, they will have higher sentence comprehension accuracy. We calculated mean percent sentence comprehension accuracy across case-marked trials, as those trials are unambiguous as to who is doing what to whom (Table 4.2).

On the first sentence comprehension block on Day 1 of the study, sentence comprehension accuracy was lower for participants learning the flexible constituent order languages than the fixed constituent order languages. This difference is especially apparent for the VI-DET languages, where the mean sentence comprehension accuracy for the VI-DET fixed constituent order language was 95.93% and the mean accuracy for the VI-DET flexible constituent order language was 68.75% with no overlap in 95% confidence intervals between the two groups of learners. In the flexible constituent order miniature languages, the only reliable cue to grammatical function assignment is the case marker. The lower sentence comprehension accuracy for participants learning flexible order languages suggests that at the time of the first sentence comprehension block, participants have not fully learned the function of the case marker.

Interestingly, in the first sentence comprehension block on Day 1, learners of the VM-DET languages, which were most similar to German, had numerically the lowest sentence comprehension accuracy. This runs counter to the prediction that increased similarity to the L1 would lead to participants comprehending the miniature language better or sooner. However, the 95% confidence intervals overlap with those of the groups learning the other miniature languages, so it is difficult to say much about this result.

By the second sentence comprehension block on Day 1, the difference in accuracy between the learners of fixed and flexible constituent order languages narrowed and accuracy improved, with all groups scoring above 89% on average. On Day 2, sentence comprehension accuracy was quite high, with all groups scoring above 93% on average. This shows that participants had a full grasp on the miniature languages going into the sentence production blocks of the experiment.

Table 4.2: Mean percent accuracy for unambiguous sentence comprehension trials (95% confidence intervals shown in brackets)

		VI-SFX		VI-DET		VM-DET	
Day	Block	Fixed	Flexible	Fixed	Flexible	Fixed	Flexible
1	1	94.38 [86.56, 98.75]	85.31 [74.69, 94.06]	95.93 [93.44, 98.44]	68.75 [55, 82.18]	88.12 [76.88, 97.19]	67.19 [55, 79.69]
	2	98.44 [97.19, 99.69]	89.69 [79.05, 97.81]	99.38 [98.43, 100]	90.94 [81.25, 97.82]	97.19 [94.37, 99.06]	93.75 [87.18, 98.75]
2	1	99.38 [98.44, 100]	96.56 [92.81, 99.06]	99.06 [97.81, 100]	95.63 [89.38, 99.38]	99.38 [98.44, 100]	93.75 [86.88, 98.44]
	2	99.69 [99.06, 100]	98.75 [96.88, 100]	100 [100, 100]	97.50 [94.05, 99.69]	98.44 [96.88, 99.69]	96.25 [92.81, 98.75]

Table 4.3: Mean percent of sentences with case errors in sentence production (95% confidence intervals shown in brackets)

	VI-SFX		VI-DET		VM-DET	
Day	Fixed	Flexible	Fixed	Flexible	Fixed	Flexible
1	0.21 [0, 0.63]	1.08 [0.21, 2.39]	0.33 [0, 1.00]	2.33 [0, 5.89]	0.44 [0, 1.08]	1.27 [0, 2.99]
2	0.42 [0, 1.04]	1.04 [0, 2.29]	0.42 [0, 1.04]	1.04 [0, 2.08]	0 [0, 0]	2.29 [0, 6.25]

Sentence Production Errors

We might expect that German speakers make more grammatical errors at sentence production in the miniature languages that are more unlike their L1. We report the percentage of sentences produced where case was marked on the subject or verb instead of the object (Table 4.3) and the percentage of sentences produced with a constituent order that was not in the input miniature language (Table 4.4). Overall, grammatical errors were quite low, with case errors occurring on only 0.91% and constituent order errors occurring on only 0.42% of sentences across all miniature language conditions and both days of the experiment. There is no clear pattern of difference in case errors between learners of the different miniature languages, and all groups had means of below 3% of sentences with case errors on both days of the experiment. Numerically, learners of the VI-SFX fixed order miniature language had the highest percentage of constituent order errors on both days of the experiment, but with both of these means being below 3%, it is difficult to say if this higher rate is meaningful.

The low proportions of grammatical errors throughout the experiment shows that for all miniature language conditions, participants were able to use the miniature language successfully by the sentence production block on the first day of the experiment.

Table 4.4: Mean percent of sentences with constituent order errors in sentence production (95% confidence intervals shown in brackets)

	VI-SFX		VI-DET		VM-DET	
Day	Fixed	Flexible	Fixed	Flexible	Fixed	Flexible
1	1.47 [0.42, 2.52]	0 [0, 0]	0 [0, 0]	0 [0, 0]	0.63 [0, 1.88]	0 [0, 0]
2	2.71 [1.25, 4.17]	0.21 [0, 0.63]	0 [0, 0]	0 [0, 0]	0 [0, 0]	0 [0, 0]

4.3.3 Sentence Production Analysis

We now turn to our main analysis, where we ask if participants use constituent ordering and case marking differently at sentence production depending on how similar the miniature language they learned is to German.

Constituent Order Use

We look at the productions of participants who learned the flexible order miniature languages only, as all participants who learned the fixed order miniature languages trivially matched the input proportion of 100% subject-before-object (S-before-O) order (VSO in the VI-SFX and VI-DET fixed languages and SVO in the VM-DET fixed language). We first compare proportion S-before-O usage across the flexible order miniature languages. Then we compare the proportion S-before-O usage to the input proportion of 50% for each miniature language.

To investigate whether German-speaking participants who learned miniature languages that differed in the amount of similarity to German used constituent order differently, we fit a generalized linear mixed effects model that predicted S-before-O usage from miniature language (sliding contrast coded; VI-DET compared to VI-SFX and VM-DET compared to VI-DET[1]), day of training (sum coded; 1=Day

[1] This coding scheme reflects our hypothesis that constituent order use depends on the similarity of the miniature language to the L1, with VM-DET being most similar to German and VI-SFX being the least.

2, -1=Day1), and their interactions. The model contained the fullest converging random effects structure (by-participant random intercept and by-participant random slope of day). There was no main effect of miniature language on proportion S-before-O use (VI-DET vs. VI-SFX contrast: $\hat{\beta}$ = 0.237, z = 0.586, p = 0.558; VM-DET vs. VI-DET contrast: $\hat{\beta}$ = 0.515, z = 1.269, p = 0.204, full model in Table A.8). There was a main effect of day of training, with participants using less S-before-O order on the second day of training ($\hat{\beta}$ = -0.226, z = -2.283, p = 0.022). The interaction between the comparison of the VI-DET and VI-SFX languages and day of training was not significant ($\hat{\beta}$ = -0.080, z = -0.330, p = 0.741). The interaction between the comparison of the VM-DET and VI-DET languages and day of training also failed to reach significance ($\hat{\beta}$ = 0.453, z = 1.862, p = 0.062). Thus, we do not find evidence that the proportion of each constituent order that participants used depended on the miniature language that they learned. With respect to constituent ordering, we expected learners of the flexible order VM-DET language (that has German-like verb-medial constituent ordering) to use more of the German default S-before-O ordering, but we did not see a significant difference in S-before-O use between learners of the VM-DET language and the VI-DET language (that did not have German-like constituent ordering).

We next asked whether participants learning the flexible order VI-SFX, VI-DET, and VM-DET languages differed from the input proportion (50%) of S-before-O use on the second day of training. We used a generalized linear mixed effects model that predicted S-before-O use from miniature language (treatment coded) and the fullest converging random effects structure (by-participant random intercept). The intercept of this model captures whether the miniature language condition coded as the reference level differs from the 50% input proportion of S-before-O use. We ran this model three times, with each miniature language condition coded as the reference level. As reported in Chapter 3, learners of the VI-SFX language matched the input proportion of VSO (41.5%, $\hat{\beta}$ = -0.506, z = -1.597, p = 0.110). Learners of the VI-DET language also matched the input proportion of VSO (44.7%, $\hat{\beta}$ = -0.326, z = -1.041, p = 0.298). However, learners of the VM-DET language used significantly

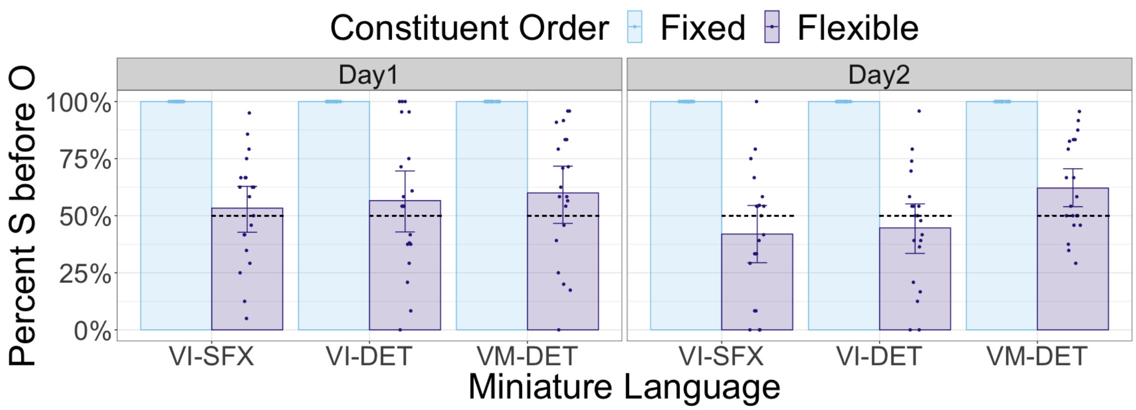

Figure 4.2: Subject-before-object constituent order use in sentence production by day of training, miniature language set, and miniature language constituent order flexibility. Y-axis is percent VSO use in the VI-SFX and VI-DET languages and percent SVO use in the VM-DET languages. The dashed line represents the 50% input percentage S-before-O order for the flexible order language (S-before-O input for the fixed order language is 100%). Dots are individual participants' percentages of S-before-O use. Error bars are bootstrapped 95% confidence intervals.

more than 50% SVO order on the final day of the study (62.1%, $\hat{\beta} = 0.630$, z = 2.022, p = 0.043). Thus, while participants learning the three flexible order miniature languages do not significantly differ from each other in amount of S-before-O order in their productions, the participants learning the VM-DET order language (the language that has the most German-like constituent ordering) use significantly more S-before-O ordering than the input of 50% on the final day of training. This shows evidence of L1 influence: participants learning the VM-DET flexible order language changed the input proportion of constituent ordering to include more sentences in the German default S-before-O order. Participants learning the languages that do not have German-like constituent ordering match the input proportion of S-before-O ordering on the final day of training, further suggesting that changing the input constituent ordering occurs because the miniature language has L1-like constituent ordering.

Case Use

We now compare how German L1 speakers who learn miniature languages with increasing similarity to German use case. Will all learners follow the bias to balance production effort against message uncertainty by using less case in a fixed order language than a flexible order language, or does this effect depend on how similar the case marker placement and constituent ordering are to the learners' L1? To answer this question, we compared the proportion of case used by participants who learned the fixed and flexible order VI-SFX, VI-DET, and VM-DET miniature languages by fitting a generalized linear mixed effects model that predicted case use from miniature language type (sliding contrast coded; VI-DET compared to VI-SFX and VM-DET compared to VI-DET), constituent order flexibility (sum coded; 1 = flexible order, -1 = fixed order), day of training (sum coded: 1 = Day 2, -1 = Day 1), and their interactions. The model contained the fullest converging random effects structure (by-participant random intercept and by-participant random slope of day). Constituent order flexibility had a main effect on case use with participants using more case in the flexible constituent order languages than in the fixed ones ($\hat{\beta}$ = 0.965, z = 4.420, $p < 0.001$, full model results in Table A.9). Learners of the VI-DET languages did not differ overall in the amount of case used from learners of the VI-SFX languages ($\hat{\beta}$ = 0.779, z = 1.458, p = 0.145). However, learners of the VM-DET languages used less case overall than learners of the VI-DET languages ($\hat{\beta}$ = -1.211, z = 2.280, p = 0.022). Recall that learners of the VM-DET flexible language used more SVO order than the input 50%, which brings the output language closer to a fixed constituent order language in that there is less uncertainty about who is doing what to whom. This could be responsible for decreased case use in the VM-DET languages compared to the VI-DET languages, where participants matched the input constituent order. There was a marginal effect of day of training ($\hat{\beta}$ = 0.213, z = 1.945, p = 0.51) with participants using somewhat more case marking on the second day of the study. There was an interaction between constituent order flexibility and the comparison of the VI-DET languages and the

VI-SFX languages ($\hat{\beta}$ = -1.521, z = -2.846, p = 0.004). Simple effects testing revealed that learners of the VI-SFX languages and VM-DET languages use less case in the fixed order language than the flexible order language (VI-SFX: $\hat{\beta}$ = 1.792, z = 4.719, p $<$ 0.001; VM-DET: $\hat{\beta}$ = 0.832, z = 2.226, p = 0.026, full model results in Table A.10), but the same did not hold for learners of the VI-DET languages. Learners of the VI-DET languages do not significantly differ in the amount of case used in the fixed and flexible order languages ($\hat{\beta}$ = 0.271, z = 0.718, p = 0.473). No other interactions reached significance (p's $>$ 0.274). In summary, participants who learned the miniature languages least like the L1 (VI-SFX) and most like the L1 (VM-DET) used more case in the flexible order miniature language than the fixed, but those who learned the miniature languages of intermediate similary (VI-DET) did not. However, we do see a numerical trend on Day 2 of VI-DET flexible constituent order miniature language learners (mean percent case-marked = 69.3%) using more case than fixed constituent order learners (mean percent case-marked = 59.7%). As this difference seems to be emerging over days of training, perhaps we would see VI-DET learners also display a significantly different amount of case use between the fixed and flexible order miniature languages if given more training. We return to this possibility in the discussion.

We were also interested in how learners of the different miniature languages differed in case use from the input of 67% case-marked sentences. In previous work investigating the bias to balance production effort against message uncertainty, participants who learned a fixed constituent order language used significantly less case than the input proportion, showing that they restructure the miniature language in a way that aligns with this cognitive bias (Fedzechkina et al. 2017). As explained in our comparison to the input analysis for constituent order use above, we used generalized linear mixed effects models that predicted case use on Day 2 from miniature language (treatment coded) and the fullest converging random effects structure (by-participant random intercept). We constructed two models, one for fixed order languages and one for flexible order languages. We ran each model three times with each miniature language condition (VI-SFX, VI-DET, VM-DET) as the reference

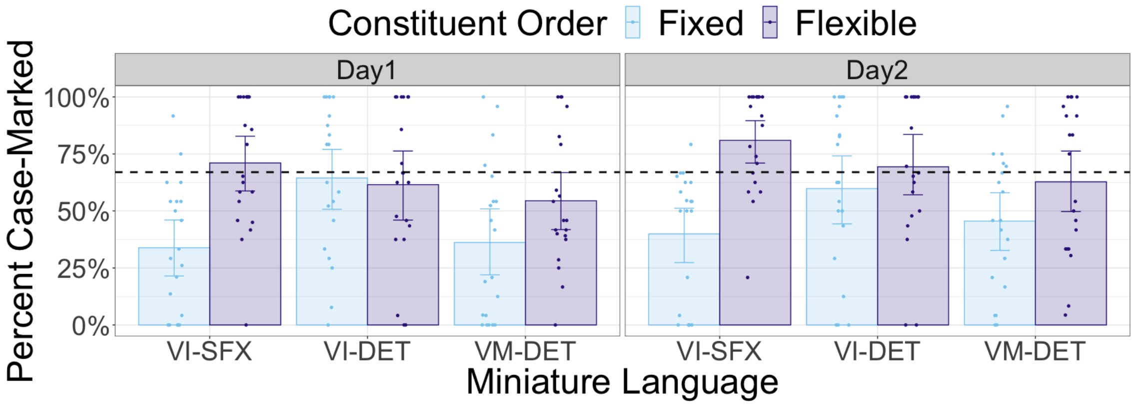

Figure 4.3: Case use in sentence production by day of training, miniature language set, and miniature language constituent order flexibility. The dashed line represents the 67% input percentage of case (same for all miniature languages). Dots are individual participants' proportions case-marked productions. Error bars are bootstrapped 95% confidence intervals.

level. For the fixed order languages, we found that participants who learned the VI-SFX and VM-DET languages used significantly less case than the input proportion (VI-SFX: $\hat{\beta}$ = -1.764, z = -3.278, p = 0.001; VM-DET: $\hat{\beta}$ = -1.147, z = -2.168, p = 0.0302) but participants who learned the VI-DET language did not significantly differ from the input proportion ($\hat{\beta}$ = -0.027, z = -0.051, p = 0.959). For the flexible order languages, we found that participants who learned the VI-SFX language used more case than the input ($\hat{\beta}$ = 2.067, z = 0.691, p = 0.003) and that participants who learned the VI-DET and VM-DET languages did not significantly differ from the input proportion of case (p's $>$ 0.100).

The results from the case marker use analysis show that earners of the miniature languages we designed to be least like German (VI-SFX) and most like German (VM-DET) follow the bias to balance production effort against message uncertainty by using significantly less case in a fixed order language than the input and than in a flexible order language, but learners of languages with intermediate similarity (VI-DET) do not, although they do use numerically more case in the flexible order language than the fixed on Day 2.

4.4 Discussion

We set out to investigate whether increasing the similarity between a miniature language and participants' L1 would cause participants to abandon a general bias to balance production effort against message uncertainty in favor of an L1-specific bias. We found mixed results. When it comes to constituent ordering, German speakers exposed to the flexible VM-DET language (the language most like German, with SVO/OVS constituent order and a determiner case marker) used more SVO than the input on the final day of training, but those exposed to the flexible VI-DET language (the language intermediately similar to German, with VSO/VOS constituent order and a determiner case marker) matched the input proportion of VSO use, as did those participants learning the VI-SFX language. In German, SVO is most often used in simple transitive sentences of the type presented in this study and OVS order is rare (Lederer 1969). Thus, the increased use of SVO in the VM-DET language does fit with an influence from German, however, it is important to note that the SVO use only significantly differed from the input proportion of 0.5 and not from the S-before-O order use of the other two miniature language conditions.

When we analyzed case marking results, we found that learners of the VI-SFX and VM-DET languages followed the bias to balance production effort against message uncertainty: using more case in the flexible order language (where case marking provides more information) than in the fixed order language. Learners of the VI-DET languages did not significantly differ in the amount of case used in the fixed and flexible order languages, however on Day 2, we see numerically higher percent case use for the flexible order VI-DET language compared to the fixed. Perhaps if the experiment included a third day of training, we would see the VI-DET language learners also follow the bias to balance production effort against message uncertainty by using significantly more case in the flexible order language than the fixed. Why would it take these learners longer to display this bias than it did for VI-SFX and VM-DET learners, for whom a difference in case use between fixed and flexible miniature languages was even present on Day 1? One possibility is that there is a U-shaped effect of similarity to the L1 on this miniature language learning task. Perhaps when the miniature languages are dissimilar to the L1, learners do not bring their L1 biases to the task and follow a general cognitive bias, and when the miniature languages are similar to the L1, learners understand the function of the components of the miniature language and follow a general cognitive bias. However, when the miniature languages share some features that are similar to the L1 and have some features that are unfamiliar to the learner, learners behave differently.

This explanation fits with the competition model of second language acquisition, which states that structures that are unique to the L2 are easier for second language learners to acquire than structures that are used in both the L1 and L2 but are instantiated differently (Tokowicz & MacWhinney 2005). Especially relevant to the current study, Tolentino & Tokowicz showed that when English L1 speakers were exposed to a miniature language version of Swedish, they could better judge the grammaticality of structures that were present in both the L1 and the miniature L2 and instantiated the same way (i.e., number agreement between demonstrative determiners and nouns) as well as structures that were only present in the miniature L2 (i.e., gender agreement) than structures that were present in both the L1 and the L2 but follow different rules (i.e., definiteness marking) (2014). Similar results have been found for classroom learners of a second language: Japanese L1 learners of Korean had slower self-paced reading rates for Korean constructions that had an L1 correspondent that surfaced differently than for Korean constructions with no Japanese correspondent (Park & Kim 2021) and German L1 learners of French had diminished ERP responses to article-noun gender mismatches in French if the gender of the noun was different than it is in German (Foucart & Frenck-Mestre 2011).

Under the competition model, we can explain our results as follows: Learners of the VI-DET languages, which share similarity to German in case marking but not in constituent order, are impeded when figuring out how grammatical function assignment works in the language system compared to learners of the VI-SFX languages which share no direct similarities to German constituent ordering and case marking and the VM-DET languages in which case marking and constituent ordering are both German-like. Thus, learners of the VI-SFX and VM-DET languages are able to realize that the case marker is a redundant cue in the fixed order language and drop it in their productions, but learners of the VI-DET languages take longer to realize this and only begin to drop case in the fixed order language after more training on Day 2. This study could be repeated with an additional day of training to see if learners of the VI-DET language significantly follow the bias to balance production effort against message uncertainty on a Day 3 sentence production task.

A remaining question is how, specifically, are learners of the VI-DET languages impeded? Sentence comprehension results do not point to VI-DET learners being unable to learn either the constituent ordering or the case marking in the miniature languages, as they could accurately select who was doing what to whom for miniature language sentences by the second comprehension test on the first day of the study (accuracy above 90% for both fixed and flexible languages). Further, learners of the VI-DET languages had low percentages of trials with case marking errors and no trials with constituent ordering errors in sentence production. Unfortunately, our paradigm focusing on language use did not allow for further investigation of learning or processing effects caused by an L1 influence. Using larger miniature languages that take more time to acquire could better allow us to compare learning differences between participants learning the VI-SFX, VI-DET, and VM-DET languages. Additionally, the effects that might explain the miniature language use differences we see are often found in eye-tracking, electroencephalography, and brain imaging studies, which allow researchers to see how L2 learners' language processing is influenced in real time (Díaz et al. 2016, Liang & Chen 2014, Wang et al. 2022, Andersson et al. 2019, Paquet 2018). Therefore, further investigation on the effects of miniature-language-to-L1 similarity could benefit from use of an on-line method. For example, employing eye-tracking during the sentence comprehension blocks could further reveal how the participants understand grammatical function assignment in the miniature language as we could see participants' eye movements as they are asked to select the scene that correctly depicts who is doing what to whom in the sentence. If our results in the current study are due to an L1 influence interfering with the acquisition of the VI-DET set of miniature languages, we might expect to see more eye movements toward the incorrect scene or more re-reading of the miniature language sentence for participants learning the VI-DET languages than participants learning the VI-SFX and VM-DET languages.

While we found some evidence of L1 influence as the direct similarity between the miniature languages and German was increased, participants learning the languages with German-like constituent order and German-like case marking still used less case in the fixed order language than the flexible one, and participants learning the languages with German-like case marking and non-German-like constituent order were trending towards doing the same on Day 2. This result, paired with the finding in Chapter 3 that speakers of structurally different L1s also used less case in a fixed constituent order miniature language than a flexible constituent order one, shows that this behavior is robust and gives more support to the idea that it comes from a general cognitive bias to balance production effort against message uncertainty.

Chapter 5

General Discussion

In this dissertation, we investigated factors that could affect miniature language learning behavior beyond the general cognitive biases that researchers claim are responsible for miniature language use patterns that mirror language universals. In Chapter 2, we investigated the effect of production modality on regularization of constituent ordering. In Chapter 3, we studied the effect of the structure of participants' L1s on constituent ordering and case marking. In Chapter 4, we asked how direct similarity between the miniature language and participants' L1 influenced constituent ordering and case marker use. The findings are summarized below.

5.1 Findings of Chapter 2

Here, we exposed participants to a miniature language with flexible constituent order: 67% VSO and 33% VOS. The participants were split into two groups, one of which typed their miniature language productions and the other of which spoke their productions into a microphone. We measured the entropy of participants' productions, which showed how much they preferred to use one constituent order over the other (in other words, if they regularized constituent order). We did not find a difference in entropy between participants who typed their productions and those who spoke their productions. However, more participants in the speaking condition failed to learn the miniature language, resulting in a higher exclusion rate. These findings, combined with the practical advantages of collecting and analyzing typed responses over spoken ones, led us to recommending using a typed production modality in miniature language learning experiments.

5.2 Findings of Chapter 3

In this chapter, we asked if participants use miniature languages with fixed or flexible constituent ordering and optional morphological case marking in the same way if their L1s make use of these cues differently. The L1s included in the study were English, which has fixed constituent ordering and no case marking on nouns, German, which has somewhat flexible constituent ordering and case marking on nouns, and Russian, which has flexible constituent ordering and case marking on nouns. We found that miniature language learners of all three L1 backgrounds use the miniature languages in the same way. They use more case marking in the flexible constituent order miniature language, where case marking gives more information as to who is doing what to whom in a sentence, than in the fixed constituent order miniature language, where the uncertainty regarding who is doing what to whom is low from constituent order alone. Regardless of how constituent order flexibility and case marking are used in the L1, miniature language learners follow a bias to balance production effort against message uncertainty. This finding extends previous results that show this pattern with English speakers alone and gives support to the claim that the bias to balance production effort against message uncertainty is a general cognitive bias that underlies the crosslinguistic pattern of more languages with fixed constituent order lacking case marking than languages with flexible constituent order.

5.3 Findings of Chapter 4

Finally, we asked if participants would exhibit an L1 influence if there were more direct similarities between the miniature language and the L1. We compared the miniature language use of German L1 speakers in Chapter 3 to that of German L1 speakers learning two new sets of miniature languages that shared increasing direct similarity to German: one set retained the verb-initial constituent ordering from Chapter 3 but used a determiner preceding the object noun as the case marker, and the most similar set used the determiner case marker as well as verb-medial constituent ordering that is present in German. We found that, like the participants in Chapter 3, participants learning the flexible order language with verb-initial constituent order matched the input 50% VSO/50% VOS constituent ordering, but participants learning the flexible order language with verb-medial constituent order exhibited an L1 influence, using more SVO (the dominant order for simple transitive sentences in German) in their productions than the input language. When it came to case marking, we found that participants learning the languages most directly similar to German followed the bias to balance message uncertainty against production effort as the participants who learned the miniature languages the least directly similar to German did in Chapter 3. However, the learners of the languages with non-German-like constituent ordering and German-like case marking matched the input proportion of case use in both the fixed and flexible order language. This effect may be explained in terms of the competition model: when presented with structures that are similar to the L1 but work differently, acquisition can be impeded. Thus, compared to learners of miniature languages that share no direct similarity to German and miniature languages that share more direct similarity to German, learners of the languages that share case marking but not constituent ordering similarity to German might not have come to the same realization about the redundancy of the case marker in the fixed order language. On Day 2, although it did not reach significance, we saw that learners of the miniature languages with intermediate similarity to German began to use more case in the flexible order miniature language than the fixed. This suggests that if participants learning these languages were given more experience with the miniature languages that were difficult to learn, they may have also followed the bias to balance production effort against message uncertainty. Although the behaviors found in this chapter could benefit from further research, the increased SVO use in the flexible order language most similar to German and the lack of difference in case use between the fixed and flexible order languages of intermediate similarity to German show that the amount of direct similarity between a miniature language and a participant's L1 can affect miniature language use.

5.4 Impact of the Findings

Together, all three studies add to the body of knowledge of best practices for implementing the miniature language learning paradigm in order to investigate language universals, and Chapters 3 and 4 offer insights into L1 influence and provide additional evidence for the bias to balance production effort against message uncertainty underlying the crosslinguistic trade-off between constituent order flexibility and presence of case marking.

5.4.1 Production Modality

Chapter 2 shows that participants are able to acquire miniature languages when they type their productions and show no difference in regularization behavior, despite the proposed learning benefits of speaking. There did not appear to be a benefit in learners hearing their own productions in learning the miniature language (cf. Baese-Berk & Goldrick (2009)). There appeared to be no difference in probability tracking between the speaking and typing modalities, even though prior research has shown a benefit in aural over visual information for statistical learning (Conway & Christiansen 2005, Saffran 2002). It is important to note that our study did include aural information (text-to-speech generated nouns and sentences) for all of the language input in both the speaking and typing conditions, so this was perhaps sufficient aural information for the typing condition. This result allows researchers to employ the more practical typing modality in their experiments with less worry that it is affecting miniature language use. In addition to this dissertation work, other studies have also used a typing modality over a speaking modality because of the results in Chapter 2 (Fedzechkina et al. 2022, Zhao & Fedzechkina 2020).

5.4.2 L1 Influence

Chapter 3 adds to the body of work that tests miniature language learning findings crosslinguistically (Culbertson et al. 2020, Martin & Culbertson 2020, Saldana et al. 2021). Especially in cases where it is not possible to find participants who do not speak a language with the features of interest for a study (e.g., where the feature is binary, as in word order harmony tested in Culbertson et al. (2020): all natural languages will either be harmonic or non-harmonic), comparing the miniature language learning preferences of speakers whose L1s differ in the features of interest can provide insight into whether or not L1 influence is at play. In our case, because speakers of languages that differ in their use of case marking and constituent order flexibility used the miniature languages in the same way, we have more evidence that participants are not dropping case in a fixed order miniature language in order to bring it closer to the structure of their L1.

In Chapter 4, we saw that L1 influence is possible in the use of miniature languages that are usually restructured to balance production effort against message uncertainty. L1 influence only manifested for learners of languages that were more similar to German than the initial languages in Chapter 3. Connecting our results to L2 acquisition, this gives support to the myriad of research that suggests that crosslinguistic influence depends on the similarity between the L1 and the target language (Hopp 2010, Rothman 2015, Sabourin et al. 2006, Slabakova 2017). Further, in the languages of intermediate similarity to German, German speakers did not significantly follow this bias. This suggests that L1 influence manifests as interference specifically when the learning of structures of the L2 (in this case, the miniature language) competes with learners' knowledge of their L1, supporting the competition model of second language acquisition (MacWhinney 2008, Tokowicz & MacWhinney 2005).

Additionally, Chapter 4 shows that how researchers construct miniature languages can result in L1 influences. As increasing direct similarity between the elements of the miniature language under study and the L1 results in different miniature language use, this study gives support to the practice of constructing miniature languages to be as different from participants' L1s as possible. Taken together, Chapters 3 and 4 offer practical advice on how to rule out L1 influence as the cause of miniature language learning behavior in studies investigating the roots of language universals: limit the similarities between the miniature language and the L1s of participants, and empirically investigate the possibility of L1 influence by comparing the performance of speakers of speakers of languages that differ in their use of the feature(s) of interest.

5.4.3 Balancing Production Effort Against Message Uncertainty

Additionally, Chapters 3 and 4 provide more evidence for the causality of the crosslinguistic trade-off between constituent order flexibility and case marking. English, German, and Russian L1 speakers learning verb-initial fixed and flexible constituent order miniature languages with an optional suffix case marker as well as German L1 speakers learning German-like verb-medial fixed and flexible constituent order miniature languages with an optional determiner case marker use more case in the flexible order language than the fixed. These results can be explained in terms of balancing production effort against message uncertainty. In both a fixed and flexible constituent order language, a case marker takes effort to produce. However, the case marker only adds a significant amount of information in the flexible order language where grammatical function assignment is more uncertain. Thus, participants use the case marker less frequently in their productions in the fixed order language where it takes effort to produce but adds little information about who is doing what to whom. Taken together with other miniature language learning results (Fedzechkina et al. 2017, Fedzechkina & Jaeger 2020, Tal & Arnon 2022), there is now a great deal of evidence that the bias to balance production effort against message uncertainty underlies the crosslinguistic tendency for many more flexible constituent order languages to employ case marking than fixed constituent order languages.

This body of work can be taken together with other miniature language learning and typological survey studies to show that languages are structured in a way that facilitates efficient communication. Take the trade-off between word length and word frequency coined by Zipf (1935) as The Law of Abbreviation: words that are used more frequently in a language are shorter in form. This pattern has also been explained in terms of balancing the competing pressures of production effort and reducing message uncertainty: The pressure to reduce production effort favors words with shorter forms, but the confusability of shorter words is higher than that of longer words in the noisy channel that is speech transmission, so the pressure to reduce message uncertainty favors words with longer forms. Thus, researchers theorize that languages have evolved to assign the most frequent meanings to the shortest word forms, as these meanings are most predictable and can be best recovered from degraded input (Ferrer i Cancho et al. 2013, Piantadosi et al. 2011a, Zipf 1949). This explanation is supported by a miniature language learning experiment that showed that participants optimize miniature language word forms in this way (Kanwal et al. 2017). Related computational modeling and experimental work shows that the distribution of phonemes and their articulation in speech also fits with the balancing of production effort and message uncertainty (Aylett & Turk 2006, King & Wedel 2020, Wedel et al. 2018, Seyfarth et al. 2016). Additional evidence that humans structure language for efficient communication comes from dependency length minimization (Fedzechkina et al. 2018, Ferrer i Cancho 2004, Futrell et al. 2015, Gildea & Jaeger 2015, Hahn et al. 2021, Hawkins 2014), semantics (Kemp et al. 2018), and discourse (Qian & Jaeger 2012). Additionally, evidence from an iterated miniature language learning study shows that languages can be restructured over generations of speakers for more efficient communication, suggesting that languages evolve to optimize the balancing of effort and message uncertainty reduction (Kirby et al. 2015). Chapters 3 and 4 of this dissertation add evidence to this literature that a bias for efficient communication is at the root of many language universals.

5.5 Areas for Future Research

The studies in this dissertation were designed to answer questions about miniature language use preferences once participants had fully acquired the miniature language. Thus, our miniature languages were relatively easy for participants to acquire and we did not see meaningful differences in comprehension and production between conditions in our studies. However, the production modality and L1 effects investigated here likely have some processing and learning components that can be seen more clearly in other set-ups that better track the acquisition of the miniature language. One straightforward follow-up would be to make the miniature languages harder to acquire, by adding more nouns and verbs to the lexicon or adding another grammatical element, such as plural agreement. Then, we could track how accuracy of miniature language production and comprehension improves over time and analyze if it differs between production modalities, amount of abstract similarity to the L1, or amount of direct similarity to the L1. Employing an online methodology such as eye-tracking or EEG could also shed light on the miniature language learning process as we investigate these factors. For instance, it has been found that participants who are given a miniature language sentence and are asked to choose a scene increase fixations to the target scene as they learn the function of the miniature language elements (Andringa 2020, Curcic et al. 2019, Wonnacott et al. 2008). We could implement our sentence comprehension test with eye-tracking and measure the proportion of fixations to the correct scene participants make over time, and see if the rates at which participants begin fixating more on the correct scene differ depending on the experimental condition.

Another limitation of Chapters 3 and 4 is that while our research questions focused on L1, participants were not all monolingual. We recruited German and Russian speakers from a crowdsourcing platform that caters to studies and surveys conducted in the English language. Thus, most of our German and Russian L1 speakers were proficient in English. This raises the question of if participants of all L1 backgrounds were using less case in the fixed order miniature language than the flexible in an attempt to bring the fixed order miniature language closer to English. However, there is no clear reason to believe that German- and Russian-speaking participants would display a crosslinguistic influence from English over one of the other non-native languages they spoke, many of which do have a case system (Table 3.2), as there was no English present in the recruitment and experiment materials for German and Russian speakers. How an L2 or L3 influences miniature language learning behavior is an open question. To further validate that the effect seen in

Chapter 3 is due to a general cognitive bias rather than an English-specific bias, the study could be repeated with monolingual German and Russian speakers.

Additionally, the research questions asked in Chapters 3 and 4 can be asked again with other miniature languages that have been designed to probe the causality of language universals, but have only been run with participants of one language background. Conducting crosslinguistic comparisons will allow researchers to pinpoint which miniature language preferences are due to L1 influences and which can be attributed to general cognitive biases that have shaped the distribution of language structures throughout the world.

5.6 Conclusion

Over the past two decades, miniature language learning has been used to try to answer a central question of Linguistics: Why are the languages of the world structured in the way they are? Researchers have used miniature language learning results to propose general cognitive biases that underlie language universals, but some factors that could also account for miniature language behavior are left as possible confounds. This dissertation takes a careful look at factors that may have an influence on miniature language learning behavior aside from general cognitive biases, provides methodological recommendations, and importantly shows that one proposed general cognitive bias–that to balance production effort against message uncertainty–is robust across structurally different L1s.

Appendix A

Full Model Results

A.1 Chapter 2

Table A.1: Results from the linear regression model predicting entropy from production modality condition in the sentence production block

Fixed Effect	Estimate	SE	t-value	p-value
(Intercept)	0.516	0.103	5.005	> **0.0001**
Typing vs. Speaking	-0.019	0.146	-0.129	0.898

Table A.2: Results from the linear regression model predicting entropy from production modality condition with the typing condition entropy compared to the input language entropy

Fixed Effect	Estimate	SE	t-value	p-value
(Intercept)	-0.421	0.103	-4.086	> **0.001**
Speaking vs. Typing	0.019	0.146	0.129	0.898

Table A.3: Results from the linear regression model predicting entropy from production modality condition with the speaking condition entropy compared to the input language entropy

Fixed Effect	Estimate	SE	t-value	p-value
(Intercept)	-0.402	0.103	-3.903	> **0.001**
Typing vs. Speaking	-0.019	0.146	-0.129	0.898

A.2 Chapter 3

Table A.4: Results from the generalized linear mixed effects model predicting VSO constituent order use from L1, day of training and their interactions.

Fixed Effect	Estimate	SE	z-score	p-value
(Intercept)	0.004	0.133	0.027	0.978
German vs. English	0.064	0.325	0.197	0.843
Russian vs. German	0.468	0.325	1.444	0.148
Day 2	-0.128	0.042	-3.096	**0.0019**
German vs. English: Day 2	-0.288	0.102	-2.814	**0.0049**
Russian vs. German: Day 2	0.097	0.100	0.956	0.334

Table A.5: Simple effects test results from the generalized linear mixed effects model predicting VSO constituent order use from L1, day of training and their interactions.

Fixed Effect	Estimate	SE	z-score	p-value
(Intercept)	0.004	0.133	0.027	0.978
German vs. English	0.064	0.325	0.197	0.844
Russian vs. German	0.469	0.325	1.444	0.149
English: Day 2	0.030	0.073	0.420	0.674
German: Day 2	-0.257	0.071	-3.619	**0.0003**
Russian: Day 2	-0.160	0.071	-2.246	**0.024**

Table A.6: Results from the generalized linear mixed effects model predicting case use from L1, miniature language constituent order flexibility, day of training and their interactions.

Fixed Effect	Estimate	SE	z-score	p-value
(Intercept)	0.359	0.202	1.774	0.076
Flexible vs. Fixed	1.286	0.202	6.344	$\mathbf{2.23 \times 10^{-10}}$
German vs. English	-0.467	0.492	-0.949	0.342
Russian vs. German	-0.729	0.491	-1.484	0.137
Day 2	0.248	0.084	2.952	**0.0031**
Flexible: German vs. English	0.465	0.493	0.945	0.344
Flexible: Russian vs. German	-0.657	0.491	-1.337	0.181
Flexible: Day 2	0.187	0.080	2.334	**0.0195**
German vs. English: Day 2	-0.257	0.183	-1.399	0.161
Russian vs. German: Day 2	-0.311	0.179	-1.730	0.083
Flexible: German vs. English: Day 2	-0.195	0.184	-1.057	0.290
Flexible: Russian vs. German: Day 2	0.029	0.179	0.163	0.870

Table A.7: Simple effects test results from the generalized linear mixed effects model predicting case use from L1, miniature language constituent order flexibility, day of training and their interactions.

Fixed Effect	Estimate	SE	z-score	p-value
(Intercept)	0.359	0.202	1.774	0.076
Flexible vs. Fixed	1.286	0.202	6.344	**2.23 x 10^{-10}**
Fixed: German vs. English	-0.933	0.689	-1.353	0.176
Flexible: German vs. English	-0.001	0.703	-0.003	0.998
Fixed: Russian vs. German	-0.70	0.694	-0.104	0.917
Flexible: Russian vs. German	-1.386	0.695	-1.994	**0.046**
Fixed: Day 2	0.060	0.113	0.539	0.590
Flexible: Day 2	0.436	0.119	3.645	**0.0002**
Fixed: German vs. English: Day 2	-0.062	0.247	-0.252	0.800
Flexible: German vs. English: Day 2	-0.452	0.272	-1.661	0.096
Fixed: Russian vs. German: Day 2	-0.340	0.251	-1.353	0.176
Flexible: Russian vs. German: Day 2	-0.281	0.256	-1.097	0.272

A.3 Chapter 4

Table A.8: Results from the generalized linear mixed effects model predicting SO constituent order use from miniature language learned, day of training and their interactions.

Fixed Effect	Estimate	SE	z-score	p-value
(Intercept)	0.161	0.165	0.977	0.329
VI-DET vs. VI-SFX	0.237	0.405	0.586	0.558
VM-DET vs. VI-DET	0.515	0.405	1.269	0.204
Day 2	-0.226	0.099	-2.28	**0.022**
VI-DET vs. VI-SFX: Day 2	-0.080	0.243	-0.330	0.741
VM-DET vs. VI-DET: Day 2	0.453	0.243	1.862	0.062

Table A.9: Results from the generalized linear mixed effects model predicting case use from miniature language learned, miniature language constituent order flexibility, day of training and their interactions.

Fixed Effect	Estimate	SE	z-score	p-value
(Intercept)	0.591	0.217	2.717	**0.0066**
Flexible vs. Fixed	0.965	0.218	4.42	$\mathbf{9.89 \times 10^{-6}}$
VI-DET vs. VI-SFX	0.779	0.534	1.458	0.145
VM-DET vs. VI-DET	-1.211	0.531	-2.280	**0.023**
Day 2	0.213	0.109	1.945	0.052
Flexible: VI-DET vs. VI-SFX	-1.521	0.534	-2.846	**0.0044**
Flexible: VM-DET vs. VI-DET	0.561	0.531	1.057	0.291
Flexible: Day 2	0.122	0.111	1.092	0.275
VI-DET vs. VI-SFX: Day 2	-0.255	0.269	-0.947	0.344
VM-DET vs. VI-DET: Day 2	0.307	0.265	1.156	0.248
Flexible: VI-DET vs. VI-SFX: Day 2	0.154	0.270	0.569	0.569
Flexible: VM-DET vs. VI-DET: Day 2	-0.282	0.264	-1.068	0.285

Table A.10: Simple effects test results from the generalized linear mixed effects model predicting case use from miniature language learned, miniature language constituent order flexibility, day of training and their interactions.

Fixed Effect	Estimate	SE	z-score	p-value
(Intercept)	0.591	0.218	2.716	**0.0066**
VI-DET vs. VI-SFX	0.779	0.534	1.458	0.145
VM-DET vs. VI-DET	-1.211	0.531	-2.280	**0.0226**
Flexible: VI-SFX	1.792	0.380	4.719	$\mathbf{2.37 \times 10^{-6}}$
Flexible: VI-DET	0.271	0.378	0.718	0.473
Flexible: VM-DET	0.832	0.374	2.226	**0.026**
VI-SFX: Day 2	0.280	0.190	1.476	0.140
VI-DET: Day 2	0.0252	0.192	0.131	0.895
VM-DET: Day 2	0.332	0.184	1.803	0.071
Flexible: VI-SFX: Day 2	0.113	0.194	0.583	0.560
Flexible: VI-DET: Day 2	-0.016	0.185	-0.084	0.933

Appendix B

Additional Figures and Tables

B.1 Chapter 2

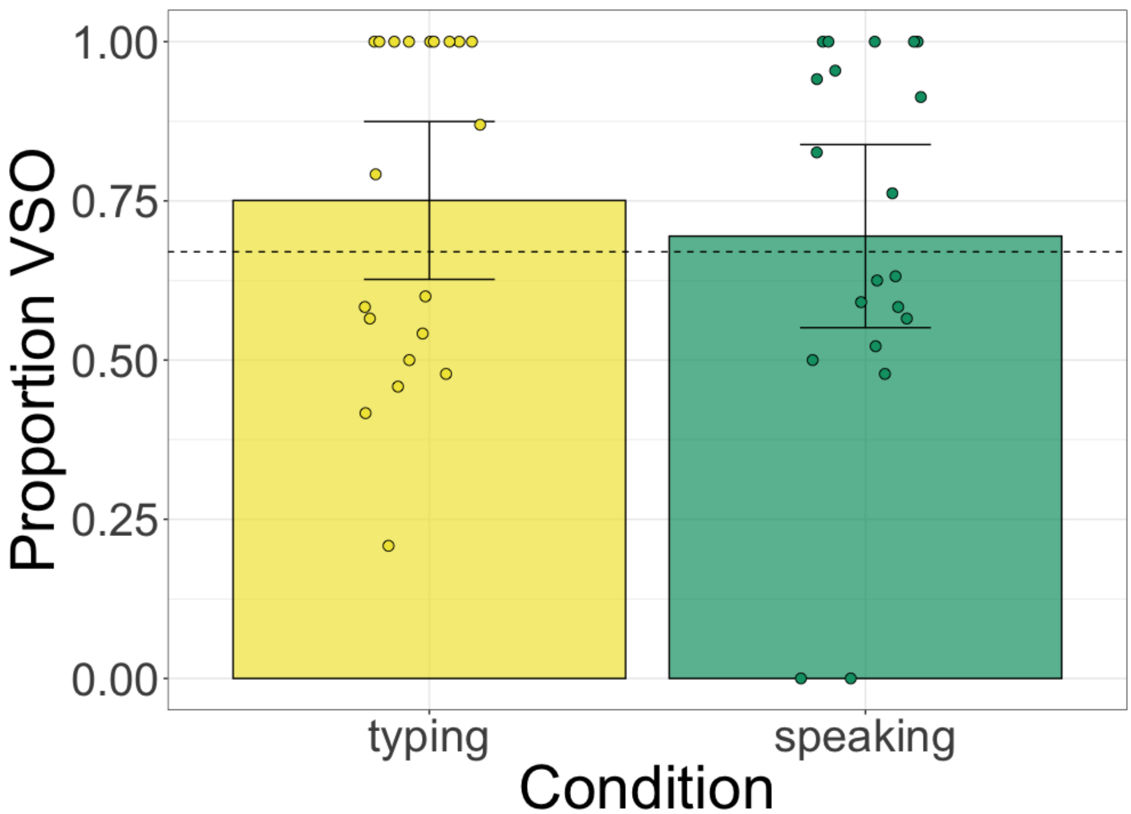

Figure B.1: Proportion VSO in participants' productions. The dashed line represents the 0.67 input language proportion. Dots represent individual participants' proportion VSO. Error bars are bootstrapped 95% confidence intervals.

Appendix C

Procedure Addenda

C.1 Chapter 3

C.1.1 Sentence Production Instructions

Now it's time for the final challenge. This time it's your job to write a sentence to say what is happening in the video. To make it easier, I'll tell you the name of the action in my language so that you can use it in your sentence.

- Please make sure to write a sentence in the alien language. (Don't just simply name the characters or action!)
- Please make sure your sentence corresponds to what is depicted in the video.
- Please do NOT use words that are not in the alien language.

Click 'Continue' to start.

References

Andersson, Annika, Susan Sayehlt & Marianne Gullberg. 2019. Language background affects online word order processing in a second language but not offline. *Bilingualism: Language and Cognition* 22(4). 802–825. https://doi.org/10.1017/S1366728918000573.

Andringa, Sible. 2020. The emergence of awareness in uninstructed l2 learning: a visual world eye tracking study. *Second Language Research* 36(3). 335–357.

Austin, Alison C., Kathryn D. Schuler, Sarah Furlong & Elissa L. Newport. 2022. Learning a Language from Inconsistent Input: Regularization in Child and Adult Learners. *Language Learning and Development* 18(3). 249–277. https://doi.org/10.1080/15475441.2021.1954927.

Aylett, M P & A Turk. 2006. Language redundancy predicts syllabic duration and the spectral characteristics of vocalic syllable nuclei. *The Journal of the Acoustical Society of America* 119. 3048. https://doi.org/10.1121/1.2188331.

Baese-Berk, Melissa & Matthew Goldrick. 2009. Mechanisms of interaction in speech production. *Language and Cognitive Processes* 24(4). 527–554. https://doi.org/10.1080/01690960802299378.

Bardel, Camilla & Ylva Falk. 2007. The role of the second language in third language acquisition: the case of germanic syntax. *Second Language Research* 23(4). 459–484. http://www.jstor.org/stable/43103748.

Bever, T G. 1970. The cognitive basis for linguistic structures. In John R Hayes (ed.), *Cognition and the development of language*, chap. 9, 279–362. John Wiley & Sons.

Bickerton, Derek. 1984. The language bioprogram hypothesis. *The Behavioral and Brain Sciences* 7. 173–221. https://doi.org/10.1017/S0305000900012459.

Bickerton, Derek. 2015. *Roots of Language*. Language Science Press.

Bivon, R. 1971. *Element Order*. Vol. 7 (Studies in the Modern Russian Language). Cambridge: Cambridge University Press.

Blake, Barry J. 2001. *Case*. Cambridge: Cambridge University Press.

Canseco-Gonzalez, Enriqueta, Laurel Brehm, Cameron A. Brick, Sarah Brown-Schmidt, Kara Fischer & Katie Wagner. 2010. Carpet or cárcel: the effect of age of acquisition and language mode on bilingual lexical access. *Language and Cognitive Processes* 25(5). 669–705. https://doi.org/10.1080/01690960903474912. https://doi.org/10.1080/01690960903474912.

Casler, Krista, Lydia Bickel & Elizabeth Hackett. 2013. Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior* 29(6). 2156–2160. https://doi.org/10.1016/j.chb.2013.05.009.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax.* Cambridge, Mass: M.I.T. Press.

Christiansen, M H. 2000. Using artificial language learning to study language evolution: Exploring the emergence of word order universals. In J L Dessalles & L Ghadakpour (eds.), *The evolution of language: 3rd international conference*, 45–48. Paris, France: Ecole Nationale Superieure des Telecommunications.

Conway, Christopher M. & Morten H. Christiansen. 2005. Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning Memory and Cognition* 31(1). 24–39. https://doi.org/10.1037/0278-7393.31.1.24.

Croft, William. 2017. Typology and Universals. In *The Handbook of Linguistics*, chap. 3, 39–55. John Wiley & Sons, Ltd. https://doi.org/https://doi.org/10.1002/9781119072256.ch3.

Culbertson, Jennifer, Julie Franck, Guillaume Braquet, Magda Barrera Navarro & Inbal Arnon. 2020. A learning bias for word order harmony: Evidence from speakers of non-harmonic languages. *Cognition* 204(November 2019). 104392. https://doi.org/10.1016/j.cognition.2020.104392.

Culbertson, Jennifer & Elissa L. Newport. 2015. Harmonic biases in child learners: In support of language universals. *Cognition* 139. 71–82. https://doi.org/10.1016/j.cognition.2015.02.007.

Culbertson, Jennifer, Paul Smolensky & Géraldine Legendre. 2012. Learning biases predict a word order universal. *Cognition* 122(3). 306–329. https://doi.org/10.1016/j.cognition.2011.10.017.

Curcic, Maja, Sible Andringa & Folkert Kuiken. 2019. The role of awareness and cognitive aptitudes in l2 predictive language processing. *Language Learning* 69. 42–71.

Currie-Hall, K, E Hume, T F Jaeger & A Wedel. 2018. The role of predictability in shaping phonological patterns. *Linguistic Vanguard*.

DeGraff, Michel. 2005. Morphology and word order in 'creolization'and beyond. *The Oxford handbook of comparative syntax*. 293–372.

Díaz, Begoña, Kepa Erdocia, Robert F. de Menezes, Jutta L. Mueller, Núria Sebastián-Gallés & Itziar Laka. 2016. Electrophysiological correlates of second-language syntactic processes are related to native and second language distance regardless of age of acquisition. *Frontiers in Psychology* 7. https://doi.org/10.3389/fpsyg.2016.00133.

Dryer, Matthew S. 1989. Large linguistic areas and language sampling. *Studies in Language* 13. 257–292. https://doi.org/10.1075/sl.13.2.03dry.

Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language*. 81–138.

Dryer, Matthew S. 2013. Prefixing vs. suffixing in inflectional morphology (v2020.3). In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Zenodo. https://doi.org/10.5281/zenodo.7385533.

der Dudenredaktion, Wissenschaftlicher Rat (ed.). 1998. *Grammatik der deutschen Gegenwartssprache*. Vol. 4 (Duden). Mannheim: Duden-Verlag.

Ettlinger, Marc, Kara Morgan-Short, Mandy Faretta-Stutenberg & Patrick C.M. M Wong. 2016. The Relationship Between Artificial and Second Language Learning. *Cognitive Science* 40(4). 822–847. https://doi.org/10.1016/j.trsl.2014.08.005.

Everett, Caleb. 2017. Languages in drier climates use fewer vowels. *Frontiers in psychology* 8. 1285.

Fedzechkina, Masha, Becky Chu & T. Florian Jaeger. 2018. Human Information Processing Shapes Language Change. *Psychological Science* 29(1). 72–82. https://doi.org/10.1177/0956797617728726.

Fedzechkina, Masha, Lucy Hall Hartley & Gareth Roberts. 2022. Social biases can lead to less communicatively efficient languages. *Language Acquisition*. 1–26. https://doi.org/10.1080/10489223.2022.2057229.

Fedzechkina, Masha & T. Florian Jaeger. 2020. Production efficiency can cause grammatical change: Learners deviate from the input to better balance efficiency against robust message transmission. *Cognition* 196. 104115. https://doi.org/10.1016/j.cognition.2019.104115.

Fedzechkina, Masha, T. Florian Jaeger & Elissa L. Newport. 2012. Language learners restructure their input to facilitate efficient communication. *Proceedings of the National Academy of Sciences* 109(44). 17897–17902. https://doi.org/10.1073/pnas.1215776109.

Fedzechkina, Masha, Elissa L. Newport & T. Florian Jaeger. 2016. Miniature artificial language learning as a complement to typological data. *The usage-based study of language learning and multilingualism*. 1–22. https://doi.org/10.1016/S0010-0277(98)00034-1.

Fedzechkina, Masha, Elissa L. Newport & T. Florian Jaeger. 2017. Balancing Effort and Information Transmission During Language Acquisition: Evidence From Word Order and Case Marking. *Cognitive Science* 41(2). 416–446. https://doi.org/10.1111/cogs.12346.

Ferdinand, Vanessa, Simon Kirby & Kenny Smith. 2019. The cognitive roots of regularization in language. *Cognition* 184. 53–68. https://doi.org/10.1016/j.cognition.2018.12.002.

Ferrer i Cancho, R. 2004. Euclidean distance between syntactically linked words. *Physical Review E* 70. 56135.

Ferrer i Cancho, R, A Hernandez-Fernandez, D Lusseau, G Agoramoorthy, M J Hsu & S Semple. 2013. Compression as a universal principle of animal behavio. *Cognitive Science* 37(8). 1565–1578.

FindingFive Team. 2019. *FindingFive: A web platform for creating, running, and managing your studies in one place*. NJ, USA. https://www.findingfive.com.

Finley, S & W Badecker. 2008. Substantive biases for vowel harmony languages. In N Abner & J Bishop (eds.), *27th west coast conference on formal linguistics*, 168–176. Somerville, MA: Cascadilla Press.

Flege, James Emil. 1989. Chinese subjects' perception of the word-final English/t/-/d/ contrast: Performance before and after training. *Journal of Acoustical Society of America* 86(5). 1684–1697.

Fodor, J D. 2001. Setting syntactic parameters. *The Handbook of contemporary syntactic theory*. 730–767.

Foucart, Alice & Cheryl Frenck-Mestre. 2011. Grammatical gender processing in l2: electrophysiological evidence of the effect of l1-l2 syntactic similarity. English. *Bilingualism* 14(3). 379–399.

Futrell, R, K Mahowald & E Gibson. 2015. Large-scale evidence of dependency length minimization in 37 languages. *Proc Natl Acad Sci USA* 112(33). 10336–10341.

Gildea, Daniel & T. Florian Jaeger. 2015. Human languages order information efficiently. *CoRR* abs/1510.02823. `http://arxiv.org/abs/1510.02823`.

Goldberg, Adele E. 2013. Substantive learning bias or an effect of familiarity? Comment on Culbertson, Smolensky, and Legendre (2012). *Cognition* 127(3). 420–426. `https://doi.org/10.1016/j.cognition.2013.02.017`.

Gomez, Rebecca L. & Louann A Gerken. 1999. Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition* 70(2). 109–135. `https://doi.org/10.1016/S0010-0277(99)00003-7`.

Greenberg, Joseph H. 1963. Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In Joseph H. Greenberg (ed.), *Universals of human language*, 73–113. Cambridge, Mass: M.I.T. Press.

Grosjean, François. 2001. *The Bilingual's Language Modes. I: Nicol, Janet L.(ed.) One Mind, Two Languages. Bilingual Language Processing. Malden.*

Hahn, Michael, Judith Degen & Richard Futrell. 2021. Modeling word and morpheme order in natural language as an efficient trade-off of memory and surprisal. *Psychological Review* 128(4). Cited by: 6; All Open Access, Green Open Access, 726–756. `https://doi.org/10.1037/rev0000269`.

Hall Hartley, Lucy & Masha Fedzechkina. 2020. Learners' bias to balance production effort against message uncertainty is independent of their native languages. In S. Denison, M. Mack, Y. Xu & B.C. Armstrong (eds.), *Proceedings of the 42nd annual conference of the cognitive science society*, 2341–2347. Cognitive Science Society.

Han, Na-rae. 2013. *The CMU Pronouncing Dictionary*. `http://www.speech.cs.cmu.edu/cgi-bin/cmudict`.

Hartmann, Frederik. 2022. Methodological Problems in Quantitative Research on Environmental Effects in Phonology. *Journal of Language Evolution* 7(1). 95–119. `https://doi.org/10.1093/jole/lzac003`.

Hawkins, John A. 1983. *Word order universals*. New York, NY: Academic Press.

Hawkins, John A. 2004. *Efficiency and Complexity in Grammar*. Oxford: Oxford University Press.

Hawkins, John A. 2014. *Cross-linguistic variation and efficiency*. Oxford, England: Oxford University Press.

Hopp, Holger. 2010. Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua* 120(4). 901–931. https://doi.org/10.1016/j.lingua.2009.06.004.

Hudson Kam, Carla L. 2019. Reconsidering Retrieval Effects on Adult Regularization of Inconsistent Variation in Language. *Language Learning and Development* 15(4). 317–337. https://doi.org/10.1080/15475441.2019.1634575.

Hudson Kam, Carla L. & Ann Chang. 2009. Investigating the cause of language regularization in adults: Memory constraints or learning effects? *Journal of Experimental Psychology: Learning, Memory, and Cognition* 35(3). 815. https://doi.org/10.1016/j.cogpsych.2009.01.001.

Hudson Kam, Carla L. & Elissa L. Newport. 2005. Regularizing Unpredictable Variation: The Roles of Adult and Child Learners in Language Formation and Change. *Language Learning and Development* 1(2). 151–195.

Hudson Kam, Carla L. & Elissa L. Newport. 2009. Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology* 59(1). 30–66. https://doi.org/10.1016/j.cogpsych.2009.01.001.

Hupp, J M, V M Sloutsky & P W Culicover. 2009. Evidence for a domain-general mechanism underlying the suffixation preference in language. *Lang Cognitive Proc* 24(6). 876–909. https://doi.org/10.1080/01690960902719267.

Iverson, Michael. 2010. Informing the age-of-acquisition debate: L3 as a litmus test. *International Review of Applied Linguistics in Language Teaching* 48. 221–243.

Jaeger, T F. 2010. Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychol* 61. 23–62. https://doi.org/10.1016/j.cogpsych.2010.02.002.

Jaeger, T F. 2013. Production preferences cannot be understood without reference to communication. *Frontiers in Psychology* 4(230). https://doi.org/10.3389/fpsyg.2013.00230.

Jaeger, T F, P Graff, W Croft & D Pontillo. 2011. Mixed effects models for genetic and areal dependencies in linguistic typology: Commentary on Atkinson. *Linguistic Typology* 15(2). 281–319. https://doi.org/10.1515/lity.2011.021.

Kanwal, Jasmeen, Kenny Smith, Jennifer Culbertson & Simon Kirby. 2017. Zipf's Law of Abbreviation and the Principle of Least Effort: Language users optimise a miniature lexicon for efficient communication. *Cognition* 165. 45–52. https://doi.org/10.1016/j.cognition.2017.05.001.

Kemp, Charles, Yang Xu & Terry Regier. 2018. Semantic Typology and Efficient Communication. *Annual Review of Linguistics* 4. 109–128. https://doi.org/10.1146/annurev-linguistics-011817-045406.

King, Adam & Andrew Wedel. 2020. Greater Early Disambiguating Information for Less-Probable Words: The Lexicon Is Shaped by Incremental Processing. *Open Mind* 4. 1–12. https://doi.org/10.1162/opmi_a_00030.

Kirby, S, M Tamariz, H Cornish & K Smith. 2015. Compression and communication in the cultural evolution of linguistic structure. *Cognition* 141. 87–102.

Kurumada, C & S Grimm. 2019. Predictability of meaning in grammatical encoding: Optional plural marking. *Cognition*.

Kurumada, C & T F Jaeger. 2015. Communicative efficiency in language production: Optional case-marking in Japanese. *Journal of Memory and Language* 83. 152–178. https://doi.org/10.1016/j.jml.2015.03.003.

Lederer, Herbert. 1969. *Reference grammar of the German language*. American e. New York: New York : Scribner.

Leech, Geoffrey N. 1975. *A communicative grammar of English*. London: London : Longman.

Li, Jia & Jim Cummins. 2019. Effect of using texting on vocabulary instruction for English learners. *Language Learning and Technology* 23(2). 43–64.

Liang, Lijuan & Baoguo Chen. 2014. Processing morphologically complex words in second-language learners: the effect of proficiency. *Acta Psychologica* 150. 69–79. https://doi.org/https://doi.org/10.1016/j.actpsy.2014.04.009..

MacWhinney, Brian. 2008. A Unified Model. In P. Robinson & N.C. Ellis (eds.), *Hanbook of Cognitive Linguistics and Second Language Acquisition*, 341–371. Routledge.

Mansfield, John, Carmen Saldana, Peter Hurst, Rachel Nordlinger, Sabine Stoll, Balthasar Bickel & Andrew Perfors. 2022. Category clustering and morphological learning. *Cognitive Science* 46(2). https://doi.org/10.1111/cogs.13107.

Marian, Viorica & Michael Spivey. 2003. Competing activation in bilingual language processing: within- and between-language competition. *Bilingualism: Language and Cognition* 6(2). 97–115. https://doi.org/10.1017/S1366728903001068.

Martin, Alexander & Jennifer Culbertson. 2020. Revisiting the Suffixing Preference: Native-Language Affixation Patterns Influence Perception of Sequences. *Psychological Science*. https://doi.org/10.1177/0956797620931108.

McAllister, R, J E Flege & T Piske. 2002. The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics* 30. 229–258.

McWhorter, John. 2022. Spoken and sign language emergence: a comparison. *Languages* 7(3). https://doi.org/10.3390/languages7030184. https://www.mdpi.com/2226-471X/7/3/184.

Michel, DeGraff. 1999. *Language creation and language change : creolization, diachrony, and development.* (Learning, Development, and Conceptual Change). A Bradford Book.

Moravcsik, Edith A. 2011. Explaining Language Universals. In Jae Jun Song (ed.), *The oxford handbook of language typology*, 69–89.

Newport, Elissa L. & Richard N Aslin. 2004. Learning at a distance I . Statistical learning of non-adjacent dependencies. 48. 127–162. https://doi.org/10.1016/S0010-0285(03)00128-2.

Pajak, B, A B Fine, D Kleinschmidt & T F Jaeger. 2016. Learning additional languages as hierarchical inference: insigts from L1 processing. *Language Learning.*

Pajak, Bozena & Roger Levy. 2014. The role of abstraction in non-native speech perception. *Journal of Phonetics* 46(1). 147–160. https://doi.org/10.1016/j.wocn.2014.07.001.

Paquet, P. 2018. Influence of l1 properties and proficiency on the acquisition of gender agreement. English. *Journal of Language and Education* 4(1). 92–104.

Park, Sun Hee & Hyunwoo Kim. 2021. Cross-linguistic influence in the second language processing of korean morphological and syntactic causative constructions. *Linguistic Approaches to Bilingualism* 12(5). 687–713.

Perfors, Andrew. 2012. When do memory limitations lead to regularization? An experimental and computational investigation. *Journal of Memory and Language* 67(4). 486–506. https://doi.org/10.1016/j.jml.2012.07.009.

Perfors, Andrew. 2016. Adult Regularization of Inconsistent Input Depends on Pragmatic Factors. *Language Learning and Development* 12(2). 138–155. https://doi.org/10.1080/15475441.2015.1052449.

Piantadosi, S T, H Tily & E Gibson. 2011a. The communicative function of ambiguity in language. *Cognition* 122. 280–291. https://doi.org/10.1016/j.cognition.2011.10.004.

Piantadosi, S T, H Tily & E Gibson. 2011b. Word lengths are optimized for efficient communication. *Proc Natl Acad Sci USA* 108(9). 3526. https://doi.org/10.1073/pnas.1012551108.

Qian, Ting & T Florian Jaeger. 2012. Cue Effectiveness in Communicatively Efficient Discourse Production. *Cognitive Sci* 36. 1312–1336. https://doi.org/10.1111/j.1551-6709.2012.01256.x.

Roberts, Gareth & Masha Fedzechkina. 2018. Social biases modulate the loss of redundant forms in the cultural evolution of language. *Cognition* 171(July 2017). 194–201. https://doi.org/10.1016/j.cognition.2017.11.005.

Rothman, J. 2011. L3 syntactic transfer selectivity and typological determinacy: The typological primacy model. *Second Language Research* 27(1). 107–127. https://doi.org/10.1177/0267658310386439.

Rothman, J. 2015. Linguistic and cognitive motivations for the Typological Primacy Model of third language (L3) transfer: Timing of Acquisition and Proficiency Considered. *Bilingualism: Language and Cognition* 18. 179–190.

Rothman, J & J Cabrelli-Amaro. 2010. What variables condition syntactic transfer? A look at the L3 initial state. *Second Language Research* 26(2). 189–218.

Sabourin, Laura, Laurie A. Stowe & Ger J. de Haan. 2006. Transfer effects in learning a second language grammatical gender system. *Second Language Research* 22(1). 1–29. http://www.jstor.org/stable/43103694.

Saffran, Jenny R. 2002. Constraints on statistical language learning. *Journal of Memory and Language* 47(1). 172–196. https://doi.org/10.1006/jmla.2001.2839.

Saffran, Jenny R., Richard N. Aslin & Elissa L. Newport. 1996. Statistical learning by 8-months-old infants. *Science* 274(5294). 1926–1928. https://doi.org/10.1126/science.274.5294.1926.

Saldana, Carmen, Yohei Oseki & Jennifer Culbertson. 2021. Cross-linguistic patterns of morpheme order reflect cognitive biases: an experimental study of case and number morphology. *Journal of Memory and Language* 118. https://doi.org/10.1016/j.jml.2020.104204.

Sandler, Wendy, Irit Meir, Svetlana Dachkovsky, Carol Padden & Mark Aronoff. 2011. The emergence of complexity in prosody and syntax. *Lingua* 121(13). New insights into the Prosody-Syntax interface: Focus, phrasing, language evolution, 2014–2033. https://doi.org/10.1016/j.lingua.2011.05.007.

Sapir, E. 1921. *Language: An introduction to the study of speech*. New York: Harcourt, Brace.

Saratsli, Dionysia, Stefan Bartell & Anna Papafragou. 2020. Cross-linguistic frequency and the learnability of semantics: artificial language learning studies of evidentiality. *Cognition* 197. 104194. https://doi.org//10.1016/j.cognition.2020.104194.

Schepens, J, F der Silk & R Van Hout. 2015. L1 and L2 distance effects in learning L3 Dutch. *Language Learning*. 1–33.

Senghas, A & M Coppola. 2001. Children creating language: How Nicaraguan Sign Language acquired a spatial grammar. *Psychological Science* 12(4). 323–328.

Senghas, A, M Coppola, E L Newport & T Supalla. 1997. Argument structure in Nicaraguan Sign Language: The emergence of grammatical devices. In E Hughes, M Hughes & A Greenhill (eds.), *Proceedings of the 21st annual boston university conference on language development*, vol. 2, 550–561. Somerville, MA: Cascadillia Press.

Seyfarth, Scott, Esteban Buz & T. Florian Jaeger. 2016. Dynamic hyperarticulation of coda voicing contrasts. *The Journal of the Acoustical Society of America* 139(2). EL31–EL37. https://doi.org/10.1121/1.4942544.

Shannon, C. E. 1948. A Mathematical Theory of Communication. *Bell System Technical Journal* 27(4). 623–656. https://doi.org/10.1002/j.1538-7305.1948.tb00917.x.

Sheehan, Michelle, Albertyna Paciorek & John N. Williams. 2019. Case/agreement matching: evidence for a cognitive bias. *Glossa* 4(1). https://doi.org/10.5334/gjgl.413.

Slabakova, Roumyana. 2017. The scalpel model of third language acquisition. *International Journal of Bilingualism* 21(6). 651–665. https://doi.org/10.1177/1367006916655413.

Smith, Kenny & Elizabeth Wonnacott. 2010. Eliminating unpredictable variation through iterated learning. *Cognition* 116(3). 444–449. https://doi.org/10.1016/j.cognition.2010.06.004.

Tal, Shira & Inbal Arnon. 2022. Redundancy can benefit learning: Evidence from word order and case marking. *Cognition* 224. https://doi.org/10.1016/j.cognition.2022.105055.

Thomason, Sarah Grey & Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. University of California Press. https://doi.org/10.1525/9780520912793.

Tokowicz, Natasha & Brian MacWhinney. 2005. Implicit and explicit measures of sensitivity to violations in second language grammar: an event-related potential investigation. *Studies in second language acquisition* 27(2). 173–204.

Tolentino, Leida C. & Natasha Tokowicz. 2014. Cross-language similarity modulates effectiveness of second language grammar instruction. *Language Learning* 64(2). 279–309.

Wade, Terence L B. 1992. *A Comprehensive Russian Grammar*. Oxford: Blackwell.

Wang, Danyang, Sarah Wang, Benjamin Zinzer, Li Sheng & Kaja Jasinska. 2022. Cross-linguistic influences of l1 on l2 morphosyntactic processing: an fnirs study. *Journal of Neurolinguistics* 63. https://doi.org//10.1016/j.jneuroling.2022.101063.

Wedel, Andrew, Noah Nelson & Rebecca Sharp. 2018. The phonetic specificity of contrastive hyperarticulation in natural speech. *Journal of Memory and Language* 100(February 2017). 61–88. https://doi.org/10.1016/j.jml.2018.01.001.

Westergaard, Marit, Natalia Mitrofanova, Roksolana Mykhaylyk & Yulia Rodina. 2017. Crosslinguistic influence in the acquisition of a third language: The Linguistic Proximity Model. *International Journal of Bilingualism* 21(6). 666–682. https://doi.org/10.1177/1367006916648859.

Wilson, C. 2006. Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science* 30(5). 945–982.

Wonnacott, Elizabeth, Elissa L. Newport & Michael K. Tanenhaus. 2008. Acquiring and processing verb argument structure: Distributional learning in a miniature language. *Cognitive Psychology* 56(3). 165–209. https://doi.org/10.1016/j.cogpsych.2007.04.002.

Zhao, Yiyun & Masha Fedzechkina. 2020. Learners' harmonic preferences are modulated by lexical retrieval difficulty. *Proceedings of the Linguistic Society of America* 5(1). 708–721. https://doi.org/10.3765/plsa.v5i1.4758.

Zipf, G K. 1935. *The Psychobiology of Language: An Introduction to Dynamic Philology*. Houghton Mifflin, Boston, MA.

Zipf, G K. 1949. *Human Behavior and the Principle of Least Effort*. Cambridge, Mass.: Addison-Wesley.