Investigating the Source of Neg-First Biases in Typology and Acquisition

by

Danielle Burgess

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Linguistics)
in the University of Michigan
2023

Doctoral Committee:

Professor Marlyse Baptista, Chair
Professor Jennifer Culbertson, University of Edinburgh
Professor Susan Gelman
Assistant Professor Savithry Namboodiripad

Danielle Burgess

dlburge@umich.edu

ORCID iD: 0000-0002-0212-5937

Acknowledgements

I have only arrived at this juncture with the support of those who believed in me, who mentored me, who led by example, and who brought joy and humor into my life.

To My Committee: Thank you all for your time, labor, feedback; and suggestions; I have grown greatly as a researcher, teacher, and person because of you. Thank you to my committee members, Jennifer Culbertson and Susan Gelman, for engaging with my work with such a careful, critical gaze and helping me to think critically about my research questions and design. To my advisors, Savithry Namboodiripad and Marlyse Baptista, who go above and beyond in their mentorship. I am so grateful to have had you as scholarly models and guides. Thank you, Savi, for pushing me to think about big questions, for modeling radical empathy, and for unveiling the norms and histories of academia in your lab group and in our one-on-one meetings. Thank you, Marlyse, for being open-minded and patient with me as I developed my project from the seed of an idea, for always advocating for me and encouraging me, and for always asking about my well-being and my family in our meetings. Thank you for arranging our advising meeting to take place at the Arboretum at the height of the 2020 COVID-19 summer.

To My Michigan Community: I owe so much to the Linguistics Department at Michigan and the friends and colleagues I have made here. Thank you to my cohort—Aliaksei Akimenka, Joy Peltier, Tzu-Yun Tung, Kelly Wright, and Jian Zhu—for your company on this journey. I am in awe of you all! Alex, I am grateful for the levity you brought to my time here and treasure your cat photos and meme edits. Joy, I am grateful to have had your company, friendship, and advice, even after we were geographically separated. Tzu-Yun, I am grateful for our quarantine era phone calls; thank you for joining me for pumpkin carving and Thanksgiving dinner. Kelly, I am grateful for your openness and encouragement. Thank you to all my fellow graduate students for the lounge conversations, writing groups, scholarly advice, career advice, board game nights, and trivia nights at Bløm, especially Dominique Canning, Yourdanis Sedarous, Lucy Chiang, Lauretta Cheng, Mathew Kramer, Justin Craft, Tamarae Hildebrandt, Dom Bouavichith, Sophia Eakins, L.R. ‘Nik’ Nikolai, Felicia Bisnath, Yushi Sugimoto, and Hayley Heaton.

In addition to my advisors' advice and mentorship, assistance and resources from this University of Michigan community made this research possible, even as the pandemic required me to substantially alter my plans. Thank you, Emily Atkinson, for helping me think through my earliest ideas for this project. Thank you, Yushi Sugimoto, for helping me translate materials for one of the experiments in this dissertation. Thank you to David Wolfe and other members of the U-M Research Software Programming staff who worked with me to adapt my experiment so that it could allow two participants to interact with each other remotely, when COVID-19 restrictions made in-person experiments impossible. Thank you to Carl Abrego, Talisha Raviere-Winston, and Jen Nguyen, for helping me carry out my research and answering my questions patiently.

To My Teachers: I am grateful to the teachers and professors I had before graduate school who created challenging and supportive learning environments, who inspired my love for linguistics and for teaching. Thank you to Herr David Gaydos and Anthony Bo for creating engaging high school classrooms. Thank you to Wayne Harbert and Miloje Despic for encouraging me to be curious about language and linguistics. Thank you to also to the rest of the Cornell Linguistics community for and for creating opportunities for undergraduates and faculty to connect.

To My Friends & Family: Thank you for supporting me wholeheartedly during this process and for giving me outlets away from my work to recenter myself when I needed it most. Thank you to my sister, Rose, for showing me examples of linguistics on social media, helping me out with some tedious data entry, and for giving great hugs. Thank you, Uncle Rick, Aunt Chris, Karina, Aidan, Gamma & Packaw Janice and Ron Bryant, and Grandma & Grandpa Marion and David Burgess for your support. Thank you, Galyne Fischer, for always being there even when we occasionally lose contact for weeks or months as we forge our own paths through life. Thank you to the entire Fischer family for taking me in, including me in your traditions, and giving me a second home.

Thank you to the friends who shared hobbies and activities with me. Thank you to Lauren Hoffman and Forged Barbell for teaching me to weightlift and be strong, inside and out. Thank you to BTS for providing a source of comfort, motivation, hope, and a caring community of friends; *I walked and somehow managed to reach the sea*. Thank you to Dominique Canning and Sam Teeple for enabling me to fly out to LA shortly after defending my dissertation prospectus to see BTS in concert. The D&D campaign I had with Tim Campbell, Steven Tsai, Will Reis,

and Jesse Vitelli didn't quite outlast my time in graduate school, but it came close. I am grateful to my college friend group: Ariana Munger, Chelsea Cheng, Tammie Siew, Joyce Gao, Tim Campbell, Steven Tsai, Will Reis, and Jon Rudolph for being a steady pillar throughout our collective growth and life changes.

To My Parents: My success is a credit to your love and support. I would first like to thank my Dad, Mike, for your constant faith in me and for instilling in me a sense of curiosity about the world and about language. Thank you for sending care packages with Wegmans brand hot sauce, driving to Michigan numerous times to visit me, watching BTS concerts with me, and keeping me up to date on what is going on outside my Michigan bubble. I would like to thank my Mom, Debra. Thank you for being a model of perseverance, passion, and laughter. I hope I am making you proud.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

1, 2, 3	first, second, third person
ASP	aspect
AUX	auxiliary
COP	copula
DAT	dative
DECL	declarative
DET	determiner
DU	dual
FORM	formal
FUT	future
IMP	imperative
IMPF	imperfective
INF	infinitive
NEG	negation
NIND	non-indicative
OBJ	object
POL	polite
PREP	preposition
PROG	progressive
PROH	prohibitive
PFX	prefix (not specified or not relevant)
PRS	present
PRTCL	particle
PST	past
SBJ	subject
SG	singular
SUSP	suspective
TOP	topic

Abstract

The tendency for negation to appear early in the sentence, dubbed the "Neg-First principle" by Horn (1989:452), has been observed in the domains of typology, language contact, and language acquisition. Based on evidence from these fields, scholars have speculated about the source and universality of Neg-First biases affecting language learning or use, which may in turn shape how linguistic systems change over time. In this dissertation, I consolidate the varied evidence and proposed explanations for the Neg-First principle and lay out the challenges and limitations of inferring the source and existence of universal cognitive constraints from typological records or acquisition evidence, due to sampling bias and uncertainty about how a purported bias should be expected to interact with other linguistic patterns and constraints. I then introduce several artificial language learning experiments to test for the presence of Neg-First biases in different tasks and populations, in order to gain a better understanding of the possible sources of Neg-First tendencies in typology and acquisition, while controlling for features of the language and the learner's input.

Through a series of three experiments, I test the presence of Neg-First biases in the early stages of second language acquisition by examining whether participants exposed to an artificial language containing both preverbal and postverbal negation demonstrate a bias to overproduce preverbal negation compared to their training input. Experiment 1, testing whether English-speaking participants demonstrate a preference consistent with a Neg-First bias, produced inconclusive results. Participants did show a small numerical tendency to produce more preverbal negation than was present in the training input, but this effect was not statistically

significant. Experiment 2 addresses potential limitations of Experiment 1 by adding a dyadic interaction component, to test the hypothesis that Neg-First biases are driven by a communicative desire to reduce potential for the listener to misunderstand the intention of the sentence and removing the presence of a word bank to encourage greater regularization by making lexical retrieval more difficult. The results of Experiment 2 do indicate a bias to overproduce preverbal negation among English speakers, consistent with a Neg-First bias. However, interaction did not boost the use of preverbal negation relative to the solo production phase, except in the Majority Preverbal Negation Condition, where the bias to regularize the majority order in the training language was consistent with a bias to use preverbal negation. Finally, Experiment 3, tests for a preverbal negation preference among speakers of a postverbal negation language (Japanese), to investigate the role of previous language experience on the preference to produce preverbal negation in adult L2 acquisition. While English speakers show a preference to produce preverbal negation even in the absence of communication, Japanese speakers do not.

These results challenge the idea of a universal preference for preverbal negation in acquisition as a plausible candidate for shaping linguistic typology, though more nuanced explanations which retain the role of universal biases at play in language processing and acquisition are viable avenues for future exploration. These possibilities underscore the importance of examining the role of language experience, development, and contact when inferring universal principles propelling language variation and change.

Chapter 1 Introduction

Negation is an appealing area of study for those interested in cross-linguistic variation, similarity, and change, and the mechanisms that determine which patterns are common and which are rare. First, negation seems to be a universal category; no languages have been found to be incapable of expressing negation. Second, negation conveys an important and deceptively straightforward meaning; from the viewpoint of propositional logic, negation changes the truth value of a proposition. However, there are various intriguing asymmetry phenomena associated with negation vis-à-vis affirmation in terms of semantics, pragmatics, and comprehension. As Miestamo (2005: 7) writes:

> From the cognitive point of view, negative sentences take longer to process and to interpret than their affirmative counterparts; from the pragmatic point of view, negative sentences are typically used in contexts where the corresponding affirmative is present as background knowledge; and from the semantic point of view, various semantic domains are reorganized and interpreted differently under negation (see e.g. Wason and Johnson-Laird 1972; Givón 1978; Horn 2001).

We also see asymmetries in the morphosyntactic expression of negation. A prominent asymmetry is that negation is overtly marked in almost all languages, in comparison to the affirmative, which is not overtly marked. The focus of this dissertation is another structural asymmetry: the apparent tendency for negation to occur before the verb in spoken languages.

Horn (1989: 452) dubbed this the Neg-First principle, the "strong tendency for negative markers to gravitate leftward so as to precede the finite verb or other possible loci of negation". Similar tendencies or principles regarding the placement of negation early in the sentence or prior to the verb have been identified or proposed by many other scholars in work on typology (Dryer 1988, Jespersen 1917), language acquisition (Pienemann 1998, Dimroth 2010), and creole

and pidgin formation (Plag 2008, Bickerton 1981). Assuming a model in which cross-linguistic statistical tendencies emerge from language patterns being preferred or dispreferred in language learning and use as a result of properties of the cognitive system which facilitate acquisition or processing of certain patterns over others, it seems attractive to explain this asymmetry as a reflection of the conceptual asymmetries of negation noted by Miestamo above. The central aim of this dissertation is to identify and test links between asymmetries in language acquisition and use of negation at the individual level and cross-linguistic asymmetries favoring the early or preverbal placement of negation.

The degree to which languages are shaped by individual language users' cognitive constraints and biases in language learning and use is a fertile area of investigation drawing evidence from corpus, experimental, and information-theoretic methods. The Neg-First principle is an exciting phenomenon to investigate because, while it has been oft-observed by scholars of typology, language contact, and language acquisition, and several proposals have been put forward to motivate this preference in terms of functional-adaptive pressures (Pienemann 1998, Horn 1989, Dryer 1988), these proposals have yet to be corroborated through direct behavioral evidence.

Acquiring behavioral evidence is important because rather than reflecting biases of cognition or communication, certain tendencies may be a result of sampling bias, merely reflecting which languages or language families happened to survive, propagate, colonize at the expense of others. As Tily & Jaeger (2011) note, "the statistical power to detect an effect [in quantitative typology] (e.g., to detect universal tendencies) is reduced by uncertainty about the genealogical relations between languages, the period of time during which a language was spoken, and the amount of contact with other languages" (497). There has therefore been a

widespread shift in the field of typology to test inferences and theories based on typological distributions with patterns of behavior in individuals.

Throughout this dissertation, I may alternate between discussing "Neg-First tendencies", "Neg-First biases", and "the Neg-First principle". These terms are meant to differentiate different levels of explanation: I try to use "Neg-First tendencies" to refer to the *explananda*, the cross-linguistic prevalence of early or preverbal negation in domains of typology and language acquisition, "Neg-First biases" to refer to behavioral asymmetries which favor early negation at the level of the individual, and "the Neg-First principle" to refer to universal cognitive constraints which are proposed to underlie individual-level Neg-First biases and ultimately cross-linguistic Neg-First tendencies.

This dissertation employs artificial-language learning to test the existence of a behavioral bias towards preverbal negation and investigate what possible cognitive pressures may be partially responsible for the cross-linguistic prevalence of preverbal negation. The goal of this research is to contribute to a better understanding of the cognitive processes and biases affecting language learning and use, which can illuminate general mechanisms of language change which shape language patterns over time. The dissertation will address the following research questions surrounding Neg-First tendencies:

a. Is there behavioral evidence for a bias to produce early negation?
b. How do pressures on learning and communication interact to drive selection of negation ordering?
c. To what extent is a bias to produce early negation shaped by previous language experience?

In the remainder of this introduction, I will first define the empirical domain, standard negation in Section 1.1, before providing a brief summary of each chapter of this dissertation.

1.1 Defining and Classifying Negation

In this dissertation, the primary focus is "standard negation," defined by Miestamo (2005: 1) as "the basic way(s) a language has for negating declarative verbal main clauses". I will also use the term "negator" to refer to the words and morphemes that express negation. In English, the sentence, *Patches did not eat*, illustrates standard negation, marked by the standard negator *not*. Other types of negation like negative imperatives (*Do not run*), negative copular constructions (*That is not food*), and lexical negation (*unhappy*) do not fall under the category of standard negation and are not the focus of the experimental work in the current dissertation. However, whether the Neg-First principle should extend to other types of negation is an interesting question that will depend on the underlying motivation(s) for Neg-First (see discussion in Chapter 2). This dissertation's focus on standard negation is largely because standard negation is both the focus of much of the previous research in typology, acquisition, and language emergence where Neg-First tendencies have been discussed, and the focus of much of the previous psycholinguistic research on processing of negation.

It will also be helpful to illustrate different morphosyntactic classifications of standard negation. Previous literature on the typology of negation distinguishes roughly three types of negation: (i) morphological negation, (ii) particle negation, and (iii) negative verbs (Dahl 2010:12). Morphological negation is often expressed as an affix on a verb or an auxiliary, as in Turkish:

(1) Turkish (Dahl 2010: 14)
a. *Oku-yor-um*
read-PROG-1SG
'I am reading.'
b. *Oku-mu-yor-um*
read-NEG-PROG-1SG
'I am not reading.'

Particle negation uses one or more uninflected negators that are free morphemes rather than affixes, as in Polish, which uses an uninflected preverbal negative particle:

(2) Polish (Paloposki 1999: 26, 116)
a. *czyta-m*
read-1SG
'I read.'
b. *nie* *czyta-m*
NEG read-1SG
'I don't read.'

Negative verbs (unlike particle negation) incorporate at least some of the inflectional categories which characterize finite verbs, as in Finnish, where *ei* in (3b) is a negative auxiliary which agrees with the subject:

(3) Finnish (Dahl 2010: 21)
a. *Pekka* *lukee*
Pekka read.PRS.3SG
'Pekka is reading.'
b. *Pekka* *ei* *lue*
Pekka NEG.3SG read
'Pekka is not reading.'

It is, of course, non-trivial to classify many negators into these types and different scholars may privilege different strategies and make different choices when the data are ambiguous. For example, it may be difficult to determine whether to treat a given bound negative marker as an affix or a clitic (see Dahl 2010 for discussion). However, these distinctions are relevant because, as will be discussed in Chapter 2, different morphosyntactic classifications of negation may be more or less likely to conform to the Neg-First principle.

Finally, we will consider the typical discourse function of negatives. There are several different semantic functions associated with negation, including rejection, nonexistence, denial, prohibition, and inability. Standard negation is primarily associated with denial of an explicit assertion or an implicit one (something expected or contextually inferred) (Tottie 1991). Example (4a) illustrates an explicit denial, where the corresponding affirmative proposition was explicitly uttered in the conversation, while example (4b) shows an implicit denial, which denies a proposition that wasn't directly asserted.

(4) a. – That dress must have been pretty expensive. (Horn 1989: 182–183)
– It wasn't (expensive), in fact I bought it on sale.
b. – There's a sale on dry cat food.
– Patches doesn't eat dry food.

Tottie (1991) observed that a majority of all uses of negatives in English conversation data functioned as denials, with the majority being implicit denials. Implicit denials are a part of cooperative communication; the sender accounts for the receiver's likely inferences and denies incorrect inferences so that the receiver is not misled. The desire to correct the receiver's possible misconceptions as soon as possible is also central to one proposed motivation for the Neg-First principle (Dryer 1988, Horn 1989).

1.2 Evidence and Proposed Motivation for Neg-First Tendencies

In Chapter 2, I will review previous observations and proposed explanations from the fields of typology, language acquisition, and pidgin and creole studies with regards to a preference for preverbal negation.

Additionally, some other factors affecting the order of negation are described in Chapter 2, as Neg-First is not the only principle argued to play a role in the synchronic and diachronic positioning of negators. Another type of principle, separate but sometimes in conflict with a Neg-

First principle, centers structural alignment with basic constituent order. For example, the expectation that negators should behave like adverbs yields a prediction, following Greenberg's (1963) Universal 7, that in OV languages, negation will appear to the left of the finite verb, and in VO languages, negation will appear to the right of the finite verb (cf. Vennemann 1974). Treating sentence negation as a sentence qualifier yields the opposite prediction: in OV languages a sentence qualifier will typically follow the verb and in VO languages it will precede the verb (cf. Harris 1978).

As stated previously, the core focus of this dissertation is standard negation. However, whether Neg-First and other principles affecting the position of negation can or should be extended to other types of negation will depend on the explanatory mechanisms involved. For example, proposed explanations of Neg-First centering the pragmatic role of negation have argued that the pressure for early negation may be even greater in prohibitive sentences (cf. Horn 1989: 449-450).

Converging evidence from typology, language emergence, and language acquisition could indicate a mechanism in which biases at the level of the individual manifest in certain patterns of language acquisition or use, which in turn shape typology. Previous scholarship in typology and pidgin and creole studies has treated a Neg-First tendencies as stemming from a universal bias, but it is currently underexplored *why* there should be some sort of bias that favors early negation, particularly a universal one. We also want to consider whether it is necessary to motivate a Neg-First preference in terms of a negation-specific bias. Are preferences that we see in acquisition that favor early negation truly related to the order of the negator or are they an incidental result of difficulty with other patterns that occur in certain negative constructions?

This dissertation aims to investigate these concerns using artificial language learning methodologies.

1.3 Connecting Typology and Individual Biases With Artificial Language Learning

This dissertation contributes to relates to broader developments and debates surrounding the explanation of typological tendencies. Due to the inherent challenges of finding controlled yet robust language samples with which to estimate the relative frequency of particular language types or linguistic features, and also because our current knowledge of the world's languages may not be representative of possible human languages, there has been a widespread shift in the field of typology to corroborate inferences and theories based on typological distribution with patterns of behavior in individuals (e.g., Culbertson 2012, Tily & Jaeger 2011). In seeking this type of evidence, scholars operate under the assumption that linguistic features which are easier to learn, or which offer advantages in processing or use are more likely to be maintained and learned faithfully (that is, without change) over the course of repeated transmission than less learnable or communicatively efficient alternatives. Artificial language learning is a particularly useful tool for corroborating behavioral biases consistent with typological tendencies because it allows for features of the language and the learner's input to be carefully controlled.

In Chapter 3, I take a first step in investigating the existence of Neg-First biases at the level of the individual by testing whether English-speakers trained on an artificial language containing both preverbal and postverbal negation demonstrate a bias to overproduce preverbal negation compared to their training input (Experiment 1). The results of this experiment are inconclusive in determining the presence of behavior consistent with a Neg-First bias among English-speaking learners. Participants did show a small numerical tendency to produce more

preverbal negation than was present in the training input, but this effect was not statistically significant.

While this could indicate that learning biases favoring early negation are not present or not strong enough to be worth considering, other possibilities are considered. Firstly, one hypothesized motivation for the placing negation earlier in the sentence involves a communicative desire to reduce potential for the listener to misunderstand the intention of the sentence (Dryer 1988, Horn 1989). Experiment 1 did not elicit production in a communicative context, and so failed to truly probe this motivation. Another possibility is that there is a true learning bias underlying the small, non-significant trend towards the shift towards greater production of preverbal negation, but that Experiment 1 was not difficult enough to encourage participants to regularize. Both considerations are addressed in the follow-up experiment reported in Chapter 4.

1.4 Does a Communicative Context Affect Neg-First Biases In Production?

Chapter 4 considers the hypothesis that the Neg-First principle is motivated by a desire on the part of the speaker to reduce potential confusion in the listener. This hypothesis, considered by both Dryer (1988) and Horn (1989), proposes that speakers may be inclined to produce cues to negation earlier in the sentences, because delaying negation increases the risk of the listener to misunderstand the sentence as the opposite of the intended meaning, creating comprehension difficulty when the negative is introduced. I refer to this as the Audience Design Hypothesis, since it contends that speakers shape their utterances (perhaps unconsciously) to minimize an addressee's potential confusion or difficulty processing a late cue to negation. Audience design is an established idea in language production, though its effects can be subtle and language users clearly do not always avoid sentences which are harder to comprehend (see

Ferreira 2019 for a review of audience design literature and mechanistic framework for audience design in production).The idea of comprehension difficulty associated with negative sentences is somewhat corroborated by psycholinguistic evidence, since positive counterparts to negative sentences are often represented in the early stage of processing, especially when contextual support is absent (Tian & Breheny 2019).

To test the Audience Design Hypothesis for the Neg-First tendency, Experiment 2 uses an artificial language learning paradigm followed by an interactive game using the newly learned language. After being trained on a novel language similarly to Experiment 1, participants first produced the language in isolation (similar to the production phase in Experiment 1) before being paired with another participant, where they exchanged sentences describing images and selecting images based on their partner's description in a director-matcher task. This experimental design allows the proportion of preverbal negation in the director-matcher task to be compared to the proportion of preverbal negation in the solo production task. If audience design is the critical factor motivating the tendency to use negation earlier in the sentence, then we would expect to see more preverbal negation in the interaction phases than in the solo production phase.

The results of Experiment 2 do indicate a bias to overproduce preverbal negation among English speakers learning a new language, consistent with a Neg-First bias. However, interaction did not boost the use of preverbal negation relative to the solo production phase, except in the Majority Preverbal Negation Condition, where the bias to regularize the majority order in the training language was consistent with a bias to use preverbal negation. This result does not support the Audience Design Hypothesis for Neg-First.

Possible alternative motivations to the Audience Design Hypothesis for Neg-First production biases are considered in light of the results of Experiment 2. I also consider potential explanations, in terms of experiment design, for why participants in Experiment 2 showed a significant bias to boost preverbal negation even in solo production, compared to the participants in Experiment 1.

1.5 What Is the Role of Previous Language Experience?

Chapter 5 introduces Experiment 3, which examines whether the Neg-First bias observed in Experiment 2 is also found among speakers of Japanese, a language in which the standard negator follows the main verb. Though Experiment 2 identified a bias in production of preverbal negation among English speakers, it is possible that this is the result of language transfer, because English is a language in which the standard negator always appears before the lexical verb. Experiment 3, a partial replication of Experiment 2 with participants that are speakers of a postverbal negation language (Japanese), seeks to investigate the role of previous language experience on the preference to produce preverbal negation in adult L2 acquisition.

The results of Experiment 3 do not reveal evidence of a Neg-First bias among Japanese-speaking participants; neither do they show evidence of a bias towards producing postverbal negation. These results are considered in comparison to the results of English-speaking participants in Experiment 2, as well as to Japanese adult L2 acquisition of Korean, a natural language analog which also contains preverbal and postverbal options for marking.

The lack of a Neg-First bias among Japanese-speaking participants in Experiment 3 challenges the idea of a straightforward, universal bias to produce preverbal negation in acquisition as a plausible candidate for shaping linguistic typology. Though one possibility is that the cross-linguistic prevalence of preverbal negation results from historical forces rather than

pressures on learning and use, more nuanced explanations which maintain the role of cognitive constraints or biases (independent of language experience) remain to be explored. For example, I consider the possibility of a Neg-First bias that exists in childhood but is reduced with substantial experience with a postverbal negation language, or that *apparent* preferences for preverbal negation are more precisely driven by preferences for other linguistic elements (e.g., particle negation).

1.6 Summary and Contributions

In Chapter, 6, I again discuss the dissertation as a whole and elaborate on the key findings, contributions, and future directions of this work. As a whole, this dissertation deeply engages with the idea of explanation in typology and language change, and contributes empirical evidence to advance the understanding of the factors influencing the linear order of negation in the world's languages and in trajectories of acquisition.

Chapter 2 Evidence and Proposed Motivations for Neg-First

Negation (especially standard negation) is a relatively well-researched domain in typology because it is present in all languages and has an important but relatively clear meaning that varies little across languages. Dahl (2010: 9) calls negation a "low-hanging fruit" for typologists because the data has been historically more accessible; grammatical descriptions of languages often provide at least basic information about negation. Furthermore, the fact that negation is a core element in all languages, but relatively difficult to comprehend and late to be produced in development has also invited a great deal of scholarship in acquisition and psycholinguistics. However, the amount of explicit attention to the linear position of the negator has been more prominent in typology and acquisition than in psycholinguistic work.

In this chapter, I consolidate evidence that has been offered for a Neg-First bias in the fields of typology, pidgin and creole studies, and acquisition, before diving into proposed explanations and motivations for the observed patterns. I will describe hypotheses which hinge on the unique function of negation, in addition to other, more general factors which may affect the linear order of the negator in ways which may conform to or contradict the Neg-First principle, like Dryer's (1988, 1992) Harmonic Branching Principle.

2.1 Evidence for Neg-First

Next, I will review varied evidence for Neg-First: the tendency for negation to be marked before the main verb. I will also review coinciding linguistic phenomena which appear to mitigate or exacerbate Neg-First tendencies and point out limitations of our existing data which obscure the conclusions we can make about the origins or explanations of these tendencies.

2.1.1 Evidence from Typology

An early forerunner in the typology of negation, Jespersen (1917: 5) noted a tendency to place negators "first, or at any rate as soon as possible, very often immediately before the particular word to be negatived (generally the verb)". Later, after Greenberg's (1963) seminal work advanced understanding of cross-linguistic tendencies and word order correlations, a great deal of scholarship on the position of the negator focused on how parameters of basic constituent order resulted in different predictions for negation placement. Lehmann (1973) and Vennemann (1974) emphasized the importance of the verb's position with respect to the object (the VO/OV distinction) as an important organizing parameter. For example, Lehmann (1973: 48) states that "sentence qualifiers" such as markers of negation, causation, and reflexive action, are most commonly placed *after* verb roots in OV languages, and *before* verb roots in VO languages. Japanese illustrates this as a prototypical OV example where the negator follows the verb in example (5) and Hebrew as a prototypical VO example in example (6)

(5) Japanese (Lehmann 1973: 52)
*yoma-**nai***
read-NEG
'He does not read.'

(6) Hebrew (Lehmann 1973: 54)
lo: *ka:ta:b*
NEG write
'He did not write.'

However, questions about the grammatical category of negation open the door to an opposite, conflicting prediction about the placement of negation. Vennemann (1974) treats negators as adverbs, rather than sentence qualifiers, and Greenberg's Universal 7 predicts that in consistent OV languages, adverbs precede the verb. Thus, Lehmann (1973) and Vennemann (1974) make different predictions with regards to the position of negators relative to the verb in

consistent OV languages: sentence qualifiers typically follow the verb, while adverbs typically precede it. This conflict arises in part from questions over the scope of negation, though the relative amount of focus on explaining changes in the development of negation in Romance languages may have also contributed to this conflict. As Schwegler (1983: 298) notes, "Vennemann's view of predicate negation as adverbial may well derive from a desire to show that the placement of negative particles in Latin and French indeed accords with the basic word-order patterns of these languages". Because Latin (an OV language) used a preverbal negator *non*, while French (a VO language) introduced the postverbal (post-finite, to be more accurate) negator *pas*. If predicate negation is treated as adverbial, then the relative position of the negators in Latin and French conform to Greenbergian expectations.

Broader typological surveys of negation came with Dahl (1979), Payne (1985), and eventually Dryer's (2013) sample of 1324 languages in the WALS Online database. In the WALS sample, 71% (N = 934) of these languages contain a preverbal negative word or prefix (as in examples (7) and (8)), coded relative to the lexical verb, whereas 46% (N = 613) contain a postverbal negative word or suffix (as in examples (9) and (10)[1]). In examples (7)-(10), the negator is bolded and the main verb is underlined.

(7) English: Patches does **not** <u>eat</u> much.

(8) Pilagá: **sa**-n-<u>čo'ot</u>-a haga' yawo-'. (Vidal 2001)
'He did not tell about the women.

(9) Kresh: Kôkó <u>ãmbá</u> gõkó **´dĩ**. (Brown 1994)
'Koko did not hit Goko.'

(10) Rao: gu <u>mə</u>-**ndə**. (Stanhope 1980)
'I am not eating.'

[1] Note that some languages with bipartite negation may contain both a preverbal and postverbal negative morpheme, which is why the percentages add to greater than 100. Examples (4)-(6) come from the examples provided in Dryer (2013).

However, numbers of actual languages having a certain property might be misleading, due to historical and geographic non-independence. When multiple languages in the same language family share a characteristic, they cannot be considered independent phenomena, since the similarity can be attributed to their shared historical relationship. Dryer (1988) focused on the word order of negation with a more geographically and genetically balanced sample of 345 languages. A summary of the distribution of negation order observed in Dryer's sample is given in Table 1. The table includes both the number of languages observed with each order and the number of language families represented by those languages. Dryer notes that the family numbers are "more meaningful than the figures based on raw language numbers, which are influenced by genetic bias" (95).

Table 1. Summary of frequency of different subtypes (Dryer 1988:98).
Number of languages (number of language families in parentheses).

SOV		SVO		VSO		VOS	
NegSOV	8 (5)	NegSVO	4 (3)	NegVSO	34 (10)	NegVOS	8 (6)
SNegOV	6 (3)	SNegVO	47 (13)	VNegSO	0 (0)	VNegOS	0 (0)
SONegV	39 (15)	SVNegO	3 (1)	VSNegO	0 (0)	VONegS	0 (0)
SOVNeg	64 (18)	SVONeg	13 (4)	VSONeg	0 (0)	VOSNeg	0 (0)

Dryer's (1988) sample shows that there is a preverbal tendency for sentential negation which is relatively independent of main constituent order, although it appears that the tendency is strongest for verb-initial languages and weakest for verb final languages. This patterning becomes a little bit clearer when we look more closely at how different types of negative morphemes pattern, and the distribution of these morpheme types across different word order patterns.

Whereas Dryer (1988) combined data from all morphological types of negation (affixes, negative auxiliaries and non-verbal negative particles) to produce the results shown in Table 1, Dryer (1992) broke out negative particles which *do not* exhibit verbal properties (such as English

not) and negatives words which *do* exhibit verbal properties, bearing all or some of the verbal inflections associated with the clause, as in Diegueño in example (11), for separate analysis.

(11) Diegueño (Langdon 1970: 183)
ʔ-*u·ya·w-x* ʔ-***əma***·*w-x*.
1-know-FUT 1-NEG-FUT
'I won't know.'

Dryer (1992) found that while non-verbal negative particles showed a tendency to be placed before the verb in both VO and OV languages, negative auxiliary verbs were more likely to precede the lexical verb in VO languages and follow the lexical verb in OV languages. This follows Dahl's (1979) prediction that in verb-final languages, negative auxiliaries should tend to follow the verb, like auxiliaries in general, rather than be subject to the preverbal tendency.

Thus, Dryer (1988) proposes that the typological distribution observed in Table 1 seems to be basically accounted for by a combination of sometimes-competing principles: the Branching Direction Principle, the Negative-Plus-VO Principle (i.e., don't interrupt the VO unit), and the Negative-Before-Verb Principle (i.e., the Neg-First principle). These proposed principles are not meant to represent inviolable constraints on languages, but general preferences.

The Branching Direction Principle states that languages tend to be consistently right- or left-branching. That is, languages tend towards either a left- or right-branching ideal, in which branching (phrasal) categories tend to *precede* non-branching (non-phrasal, lexical) categories, or branching categories tend to *follow* non-branching categories. Assuming that auxiliary verbs combine with a phrasal category like a verb phrase, the Branching Direction Principle captures the tendency for auxiliary verbs to precede the verb in VO languages and follow the verb in OV languages.[2] It also captures the tendencies for VO languages to have prepositions and Noun-

[2] Dryer (2009) revisits implications of Branching Direction Theory if one assumes much flatter constituent structures. If we do not assume that the auxiliary verb combines with a phrasal category, then the Branching Direction Principle no longer predicts that auxiliary verbs precede the main verb in VO languages and follow in OV

Genitive order, and OV languages to have postpositions and Genitive-Noun order, among other patterns (see Dryer 1992, 2009 for further details).

On the other hand, Dryer's Negative-Before-Verb (i.e, Neg-First) Principle primarily explains the relative frequency of SONegV languages compared to SVNegO languages. In VO languages, the Neg-First Principle is not in conflict with the Branching Direction Principle, but in OV languages it is, so we see more balanced variation between pre- and post-verbal negation. In verb-final languages, negative auxiliaries tend towards postverbal placement, whereas negative particles tend towards preverbal placement (Dryer 1992).

Dahl (2010) notes some important remaining questions and considerations in light of the observed typological distribution of negation word order. First, OV languages tend to have more bound negators than free negation particles, whereas VO languages have more free negators. Furthermore, among affixal negation, suffixal negation appears to be far more common than prefixal negation (Miestamo 2005). These factors contribute to the variation in negation order we observe among verb-final languages. However, Dahl also notes that morphological negation is more often prefixal than comparable inflectional categories like tense and aspect, a fact which favors positing an independent Neg-First tendency.

Second, Dahl points out that the question of whether the auxiliary verb or the main verb is relevant to principles concerning the placement of negators has largely been ignored in the literature. Dahl (1979) proposed that uninflected negators tended to be placed before the finite element (i.e., an auxiliary when present or else the finite verb) and as close as possible to it. Horn (1989) also specifies the finite verb as the relevant point of reference in defining the Neg-First

languages. However, Hawkins' domain minimization principles (Hawkins 1994, 2004) and/or grammaticalization processes could also work towards explaining this typological pattern.

principle. However, Dryer (1988) and Dryer (2013) both record the position of standard negation relative to the main verb rather than the finite verb[3], sometimes obscuring the fact that the position of the negative morpheme is defined, according to the grammatical rules of the language, relative to the finite verb. For example, in German, the negator always follows the finite verb in main clauses, but the position of the negator relative to the main verb depends on whether an auxiliary verb is present (as in 12a) or absent (as in 12b).

(12) German
a. Laura hat **nicht** geschlafen.
Laura has-3SG NEG slept
'Laura didn't sleep.'
b. *Laura schläft **nicht**.*
Laura sleep-3SG NEG
'Laura doesn't sleep.'

However, a typological study by Van Olmen (2021) indicates that the choice of auxiliary verb or main verb as the reference point is not that important for capturing the general typological picture of the Neg-First tendency, since in the large majority of negation constructions, the finite verb and main verb coincide.

In this section, I have discussed the typology of the placement of negation. In summary, even once factors of geographical and historical relatedness are controlled for, there seems to be a preference for negation to precede the verb across the world's languages, especially when the negator takes the form of a particle. When the negator takes the form of an auxiliary verb, we see a stronger tendency for the auxiliary to precede the verb in VO languages and follow the verb in OV languages. In Section 2.2.3, I will discuss the ways in which diachronic processes of

[3] Dryer (2013) motivates this decision in several ways: 1) to maintain consistency with other chapters of the WALS corpus which refer to main verbs but not auxiliaries, 2) in languages with little verb morphology, it is unclear what is the finite verb, or whether a word coding tense and/or aspect should be considered an auxiliary verb, 3) when the negative itself is an auxiliary verb, it is more consistent to code its position relative to the nonfinite main verb

grammaticalization may explain some of these observations, but first I will briefly discuss the ways in which these typological observations apply to pidgin and creole languages, since preverbal negation has historically been treated as a notable feature of these languages formed under conditions of high language contact.

2.1.2 Evidence from Language Contact and Emergence

Though we've seen that preverbal negation is very common cross-linguistically, preverbal negation has been noted as a characteristic feature of (especially, Caribbean) pidgins and creoles (Bickerton 1975: 43, Holm 1988: 171, Schneider 2000). Speaking about English-related pidgins and creoles, Schneider (2000: 211) declares that the "uninflected negator *no*, placed preverbally, [...] is a feature which is practically universal".

Indeed, the Atlas of Pidgin and Creole Structures (APiCS) database (Michaelis et al. 2013) shows that for pidgins and creoles, the most common placement by far for standard negation is before the verb. Table 2, replicated from Haspelmath and the APiCS Consortium (2013), provides information about the position of standard negation in 76 pidgins and creoles. Where a language had several different negative markers with different positions, or a negative marker that could appear in several constructions, multiple values were selected for an individual language. Like WALS, the APiCS database only reports on the position of negation with respect to the main verb, ignoring its position relative to auxiliary verbs and/or tense, mood, and aspect (TMA) markers (note, too, that the distinction between auxiliary verbs and TMA markers, if there is one, is difficult to assess).

Table 2. Position of standard negation of pidgins and creoles in Atlas of Pidgin and Creole Language Structures

Feature	exclusive[4]	shared	all
Before the verb	57	8	65
Immediately after the verb	2	5	7
After verb plus postverbal object	5	2	7
Bipartite, before verb and immediately after	0	1	1
Bipartite, before verb and after object	3	1	4
Bipartite, other possibilities	1	0	1
	Representation:		76

Indeed, the tendency for negative markers to precede the verb seems quite strong in the APiCS languages. We observe that 86% of the pidgins and creoles in APiCS utilize a preverbal negative morpheme, as in example (13), whereas 20% have a construction where negation is only signaled after the verb, as in example (14).

(13) Tok Pisin (Smith & Siegel 2015)
Tupela ***no*** *slip gut long nait*.
3DU NEG sleep good PREP night
'They did not sleep well at night.'

(14) Principense (Maurer 2009)
Amanhan n sa kume pêxi ***fa.***
tomorrow 1SG FUT eat fish NEG
'Tomorrow I won't eat fish.'

But we run into the same methodological issue of non-independence that concerns all language typology: is preverbal negation common in pidgin and creole languages because of some universal tendency, or because of factors of language inheritance? A pre-verbal negation pattern involves some deviation from English and French lexifiers, but is consistent with Spanish

[4] In contrast to WALS, APiCS allows for languages to show multiple values for a feature. This table shows the number of languages which have only one feature in the *exclusive* column, those which have multiple values in the *shared* column, and the row sum in the *all* column.

and Portuguese. But this pattern has also been attributed to substrate transfer, from for example, West African languages like Yoruba and Gbe (Holm 1988: 172; Todd 1991: 21). Example (15) shows sentential negation in Gungbe, which occurs in a fixed position to the left of the future marker and the verb.

(15) Gungbe (Aboh 2010)
Kɔ̀jó **má** *ná* *xɔ̀* *kátìkát* *lɔ́*
Kojo NEG FUT buy kite DET
'Kojo will not buy the kite.'

Note also that the APiCS sample tends towards other features which are correlated with preverbal negation among languages more broadly, as discussed in the previous section: 71 of the 76 APiCS languages use SVO constituent order (the exclusive constituent order type in 61 languages) and 68 express negation with a negative particle (the exclusive negation type in 53 languages).

Thus, the evidence is muddled as to whether it represents influence from the source languages (substrate and/or superstrate, depending on the Creole language and conditions of language contact), or universal tendencies for preverbal negation which play a strong role in the high contact conditions of pidgin and creole formation, or whether it is the result of multiple causation. Before turning to proposed explanations of the typological observations so far, it will be useful to discuss the major findings related to first- and second-language acquisition of negation order.

2.1.3 Evidence from First and Second Language Acquisition

In the domain of first and second language acquisition, there is some evidence for a preverbal negation preference in the early stages of learning a language. In both first and second language acquisition, there is evidence that learners sometimes go through stages in which they

systematically produce negative sentences such that the negative marker appears earlier in the sentence than it would in adult-like or target-like speech, especially when negation order is to some extent variable in the target language.

For example, many longitudinal studies of child acquisition of English and German demonstrate that children progress through the developmental stages listed in Table 3.

Table 3. Developmental stages of sentence negation in English and German

	English	German
Stage I (Neg + S)[5]	**no** the sun shining	**nein** schaffe ich (I can't manage)
Stage II (*No/Nein* + VP)	I **no** reach it	ich **nein** schlafen (I don't sleep)
Stage III (*not/nicht* + V; *don't* + V; V + *nicht*)	Kathryn **not** go over there I **don't** go sleep	*Eric **nicht** schlafen* (Eric doesn't sleep) *Henning brauch **nicht** Uni* (Henning doesn't have to go to the university)

Table 3 shows that children acquiring English and German will often progress through a stage with sentence external positioning of negation, followed by a stage where a negator precedes the verb phrase, before progressing to adult-like negation constructions. However, it seems that this trajectory reflects difficulty with the acquisition of finiteness rather than necessarily problems with childrens' placement of negation, because as children begin using auxiliaries in their speech, their positioning of negation becomes adult-like (Dimroth 2010). In both German and English, the sentential negator (*not* or *nicht*) follows the finite verb and precedes the non-finite verb, which is in sentence-final position in German (see 16a). Whereas in English, the finite verb in negative sentences must be an auxiliary verb, leading to *do*-support, in

[5] Also notable is the use of the anaphoric negator (English: *no*; German: *nein*) in contrast to a sentential negator (English: *not*, German: *nicht).* Anaphoric negation relates to the content of an earlier sentence (*This is blue. No, it is red.)* whereas sentential negation reverses the truth value of the utterance in which it occurs (*This is not blue.*). The overuse of the anaphoric negator in children's speech can make it challenging to determine whether children are using standard negation, anaphoric negation, or even metalinguistic negation, which expresses a judgment that an utterance is in some way *inappropriate*, rather than negating the truth of a proposition. Studies by Feiman et al. (2017) and Austin et al. (2014) demonstrate that English-speaking children do not understand logical negation until after age 2. This corresponds roughly to the time when English-speaking children begin producing the word "not".

German negative sentences with no auxiliary verb, the main verb is raised from sentence-final position to the second position (V2), leaving the negator behind, as in (16b).

(16) German
a. *Laura hat **nicht** geschlafen.*
Laura has-3SG NEG slept
'Laura didn't sleep.'
b. *Laura schläft **nicht**.*
Laura sleep-3SG NEG
'Laura doesn't sleep.'

Once children begin mastering the use of auxiliaries in both English and German (Stage III of Table 3), these auxiliaries appear in finite position, preceding the negator, as shown in the utterance "I don't go sleep". Thus, the development of negation placement in these languages interacts with the development of auxiliary verb use and positioning. In contrast, in a language like Japanese, where negation is always postverbal, children never go through a stage of preverbal or pre-predicate negation placement, though they may fail to inflect the predicate in an adult-like way in early development (Kanagy 1994). Example (17a) and (17b) show an informal and formal style of Japanese standard negation, respectively; both constructions mark negation postverbally. Example (17c) shows a typical developmental error in which the verb stem is not inflected in an adult-like manner and *nai* is not analyzed as separate negation and tense components. However, errors involving production of preverbal negation is not found among children acquiring Japanese.

(17) Japanese (Kanagy 1994: 270, 258)
a. *tabe-**na**-i*
eat-NEG-NONP
'(I) don't eat'
b. *tabe-**mase-n**.*
eat-FORM-NEG
'(I) don't eat'
c. * *taberu-**nai***
eat.NONP-NEG
'(I) eat not'

Therefore, it seems like there must some element of variability as to ordering of the negator relative to other elements in the sentence for an apparent preference for preverbal negation to be evident in children's language production. Though this variability may be determined by social and grammatical rules, these conditions may be opaque to early learners.

Korean is an interesting case study when looking at variability, because it has two ways of forming negation. In 'long-form negation', as in (18a), negation follows the main verb and requires the auxiliary verb *ha* to support tense and other verbal inflections, similarly to English *do*-support, whereas in 'short-form negation', as in (18b), the negator *an* appears before the verb. Korean speakers generally deem these two constructions to be semantically equivalent, although for some speakers, post-verbal negation allows for an additional wide scope interpretation which the preverbal construction does not (Park 1998).

(18) Korean (Chang 1996: 101, 103)

a. *Yong-i TV-lul po-ci **an**-ha-yo.*
Yong-SBJ TV-OBJ see-SUSP NEG-AUX-POL
'Yong doesn't watch TV'

b. *Yong-un TV-lul **an** pwa-yo.*
Yong-TOP TV-OBJ NEG see-POL
'Yong doesn't watch TV.'

There are additionally some distributional differences between the two constructions. The literature on Korean negation frequently suggests that short-form negation is more frequent in colloquial or spoken contexts. On the other hand, Lee's (2012) study of sentential negation in the Sejong Spoken Corpus, which is made up of spoken data from interviews, tv show sitcoms, and talk show programs and news, found that long-form negation was more common overall. Lee (2012) also found that short-form negation was less frequently applied to Sino-Korean predicates and adjectival verbs containing at least three syllables.

In child acquisition of Korean, we find that children do produce preverbal negation before they produce post-verbal negation. In child acquisition of Korean, children use the preverbal negation construction at age 1;7, but do not use the postverbal negation construction until around age 3;5 (Kim 1997, Park 1998). Furthermore, Korean children will sometimes place the negator even earlier than its immediately preverbal position in adult speech, putting it before an object or adverb, as illustrated in example (19), which shows an utterance from a Korean child age 2;3 compared with the correct adult form.

(19) ***an** *sok* *sanghay* (Kim 1997: 377)
NEG inside/heart get.hurt:DECL
'(You) do not feel bad.'
(cf. *sok an sanghay*)

The same general trajectories of negation development that are observed in first language acquisition are also observed in second language acquisition (SLA). Jeon (2014) reports that L2 learners of Korean at a US university who participated in elicited imitation task had more difficulty repeating postverbal negation than preverbal negation. Trajectories of preverbal placement of negation are also found even where both the source and target language would have postverbal negation. For example, Hyltenstam (1977) demonstrated that learners of Swedish produce preverbal negation with verbs in main clauses even when this is erroneous, including L1 Turkish learners, even though Turkish expresses negation as a suffix on the main verb. Similarly, in L2 acquisition of Korean, both L1 English and L1 Japanese beginner and intermediate learners favored the preverbal (short-form) negation construction, producing significantly less post-verbal (long-form) negation compared to Korean adults, showing a general developmental sequence favoring preverbal negation, even though Japanese marks negation postverbally (Kim & Yun 2013).

Preferences to produce preverbal negation even among learners who already speak a postverbal negation language suggest that this phenomenon is not driven by transfer from knowledge of a previous language alone. However, it is not necessarily clear that trajectories which favor preverbal negation are actually related to negation order, as they might be caused by other differences that differentiate preverbal constructions and postverbal constructions, such as finite verb movement, differences in morphosyntax and inflection, or differences in the input frequencies that learners are exposed to.

Finally, trajectories of learning which favor preverbal negation depend on some level of variability in L2 acquisition, as in L1 acquisition; only when the word order is more variable do we observe that learners go through this preverbal negation stage. Adult L2 learners of Japanese, like children learning Japanese as their L1, do not go through a stage of preverbal negation (Kanagy 1994).

In summary, we see evidence from both first and second language acquisition of multiple languages that preverbal negation seems to be favored, particularly when the positioning of negation is variable in the target language.

2.2 Proposed Explanations and Motivations for Neg-First

Now having established the evidence for a preverbal preference across languages of the world (among languages broadly and among the creole and pidgin languages in the APiCS database) and in first and second language acquisition, I will review several proposals put forward to motivate *why* Neg-First tendencies seem to be prevalent in these domains.

2.2.1 Processing and Production Considerations

The prevalence of preverbal negation in creole languages has encouraged creolists to consider how universal tendencies may play a role in language contact; universal constraints and pressures on processing and production have been most explicitly considered. Debate has emerged about how these pressures would play out in scenarios of language contact to affect language change, and how or whether these pressures differ in their nature or strength in various linguistic ecologies.

For example, Schneider distinguishes between *diffusion*, the regular, internal, and largely stable transmission of linguistic features from one generation of speakers to the next, and *selection*, the outcome of a process of individual features competing with each other in language contact situations, in which one of the options is ultimately "selected" and the other(s) usually discarded (cf. Mufwene's framework of selection and competition; Mufwene 1996, 2001, 2008). Schneider (2000:206) comments in a footnote that "[t]he distinction between diffusion and selection is clearly related to superstrate and substrate effects, respectively, in creole studies, but in the present context I am not, at least not predominantly or exclusively, talking about creole languages, and it remains to be seen whether there are differences in kind or intensity between evolutionary patterns of creoles and other languages."

Schneider suggests that the prevalence of preverbal *no* in English-lexified creoles is a result of it being favored for *selection* among competing forms, due to substrate similarity and as a reflection a "natural universal tendency". Schneider does not necessarily offer *why* preverbal particle negation should be a natural universal tendency, beyond reviewing the typological and acquisition evidence already discussed above. However, Processability Theory, invoked by Plag (2008), does offer some framework for why this might be favored in selection.

Plag (2008) utilizes Pienemann's (1998, 2005) Processability Theory framework to argue that that the prevalence of preverbal negation in creoles is the result of convergence (i.e., multiple causation) between 'universal tendencies' (processability constraints on acquisition within Processability Theory) and source language influence (particularly, he argues, substrate language influence). The Processability Theory framework proposes a developmental hierarchy in syntax and morphology for all languages, based on the type of information exchange possible at each stage (formalized in terms of specific processing requirements based on Levelt's (1989) approach to language production). A simplified version of the processability hierarchy is given in Figure 1, where the first column names the example constituent structures and the second column specifies the type of information exchange possible at each stage. Figure 1 illustrates the information exchange required for the insertion of English morphemes in the Processability Hierarchy account. Table 4 shows how Processability Theory predicts the developmental trajectory of negation in English, building up from category procedures to sentence procedures. Because preverbal particle negation does not require inter-phrasal information exchange, it is predicted to be simpler to process and achieved earlier in acquisition.

While only examples from English are given here, the processability hierarchy in principle applies to the transfer of grammatical information (and therefore the predicted developmental trajectory) in any language. For example, Jeon (2014) demonstrated that learners of Korean as a foreign language also demonstrated an implicational hierarchy of negation development predicted by Processability Theory, with the most common and persistent error made by learners in an elicited imitation task being substitution of the Korean postverbal negation construction (refer to example (12)) with the preverbal construction.

Figure 1. A simplified account of the Processability Hierarchy
(from Pienemann 2015: 128)

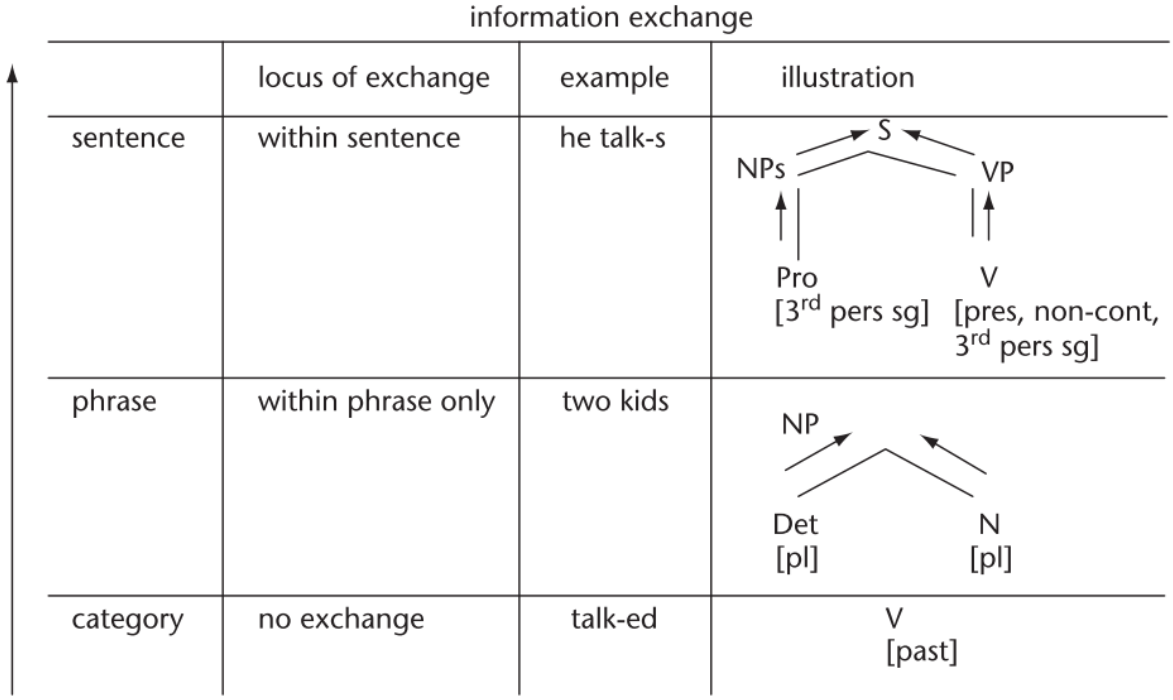


Table 4. Developmental sequence of negation for L2 English
(Larsen-Freeman & Long 1994:94, Pienemann 2000:111)

Stages	Examples	Processing Procedure
1. Clause-external	*No this one* *Not you playing here*	category
2. Clause-internal, pre-verbal	*Juana no have job* *You not go*	phrasal
3. Auxiliary + Negation	*I can't play the guitar*	sentence
4. Analyzed *don't*	*She doesn't drink alcohol*	sentence

However, beyond the application of the Processability Framework, Plag's (2008) argument makes assumptions about language contact in creole emergence scenarios which have been criticized. Plag argues that creole languages originate as conventionalized interlanguages, emphasizing that this outcome is related to the socio-historical circumstances surrounding the formation of many pidgins and creoles. Plag notes that incipient pidgin and creole speakers, especially in plantation contexts, often did not have direct access to the superstrate language and

may have fostered negative attitudes towards the superstrate language (or its speakers), negatively affecting motivation to achieve a higher degree of approximation to the superstrate language.

However, this interlanguage hypothesis has been criticized on the basis that patterns from early interlanguages would only account for a subset of the linguistic ecology at the time of emergence of a given language, given the diversity of superstrate access and heterogenous interlanguage stages in any language contact situation (Aboh 2015, Baptista 2016, Aboh & DeGraff 2017). It has also received criticism for exceptionalizing Creole formation beyond general patterns of language change. To illustrate this point, Aboh & DeGraff (2017) point out that English lost inflectional morphology, understood as a reflex of language contact phenomena (cf. Kroch et al. 2000) and loss of V-to-T raising created conditions where negation precedes the lexical verb (cf. Modern English *I do not know*, Middle English *It serveth not*). They argue that the diachronic path from French to Haitian Creole in the context of language contact with both French and Niger-Congo languages such as Gbe mirrors that of English.

Without relying on Plag's specific assumptions about the role of interlanguage and the primary agents of language change, general constraints on processing and production may be relevant to language change, especially in scenarios of language contact, more broadly.

Though Processability Theory may explain preferences for particle negation over morphological negation in terms of the information-exchange hierarchy, this theory does not seem to make any predictions about a preference for preverbal negation over postverbal negation. However, MacDonald's (2013) Production-Distribution-Comprehension (PDC) account provides another framework which aims to explain typological regularities in terms of production processes and pressures on utterance planning, execution, and memory. The PDC

account proposes three memory-related production biases which can affect the distribution of syntactic forms in languages: Easy First, Plan Reuse, and Reduce Interference. The Easy First bias has the clearest connection to the Neg-First principle; Easy First dictates that people should produce more easily planned or accessed elements first, leaving more time for the planning of more difficult elements. There are many different forces which have been argued to affect ease of planning or access in production such as frequency (more frequent words are produced earlier), length (shorter words and phrases are produced earlier), conceptual salience (elements more important to the producer or which the producer is attending to due to task-specific reasons are produced earlier), and givenness (information previously given in the discourse is produced earlier). A general preference for preverbal negation might be maintained under the Easy First bias in the PDC framework if, perhaps, negation is considered to be a conceptually salient or important element in the production planning process.

2.2.2 Communicative Considerations

Another hypothesis is that the Neg-First Principle is motivated by communicative considerations of audience design. Recall that Dryer (1988) proposed a combination of three sometimes-competing principles account for typological distribution observed in Table 1: the Branching Direction Principle, the Negative-Plus-VO Principle (i.e., don't interrupt the VO unit), and the Negative-Before-Verb Principle. Dryer (1988: 102) provides the following functional motivation for the Negative-Before-Verb Principle:

> Negative morphemes carry a large communicative load in the sense that they carry an important part of the message. If a hearer fails to hear the negative morpheme in a sentence, they will have fundamentally misunderstood the sentence. […] Delaying them increases the risk of misunderstanding, creating a kind of 'semantic garden path', since the apparent meaning of a sentence up to but not including the negative will be the opposite of the intended meaning.

Dryer's 'semantic garden path' explanation seems grounded in the assumption that speakers are constructing or activating the corresponding affirmative sentence at least until negation is cued.

This proposal appears to be supported by various behavioral and neurolinguistic studies which seem to indicate that the affirmative sentence is initially considered and represented during the processing of negation. Numerous experimental studies investigating the processing of negation (Carpenter & Just 1975; Clark & Chase 1972; Kaup et al. 2006, 2007) found that English speakers are slower and make more errors when responding to negative sentences than their affirmative counterparts. This general finding on the delayed interpretation of negation has been explained by a two-step theory of negation processing (MacDonald & Just 1989). This theory proposes that a negative sentence initially facilitates the representation of a situation described by its affirmative counterpart and then shifts to the representation of a negated state. If processing occurs this way across languages and contexts, then this could provide a good motivation for speakers to avoid delaying negation, as Dryer (1988) suggests.

However, more recent research has found that with an appropriate context, listeners do not process negation in two steps but immediately obtain the negative interpretation (Nieuwland & Kuperberg 2008; Tian et al. 2010). Tian et al. (2010), for example, examined how participants comprehended simple negative sentences (e.g., *Mike didn't iron his shirt*) and cleft sentences with a negative clause (e.g., *It was Mike who didn't iron his shirt*); because clefts are known to have a presupposition (e.g., that someone didn't iron his shirt), they create a pragmatic context for negative meaning. Tian et al. (2010) found that after reading simple negative sentences, participants responded faster to a picture of the affirmative situation (e.g., an ironed shirt) than to a picture of the negative situation (e.g., a crumpled shirt), in line with the delayed interpretation

found in previous studies when context is absent. On the other hand, after reading cleft negative sentences, participants responded faster to the picture of the negative situation.

Such results suggest that earlier findings of delayed processing of negation reflect the infelicitous use of negation. This is because, as Wason (1965, 1972) argued, the negation requires a felicitous context that triggers the listener's anticipation of its use. For example, with the presence of an apple in a naturalistic situation, there is no pragmatic reason to say "This is not a pear," even if it is true. In contrast, in a situation where several pears and an apple are present, it is more plausible to say "This is not a pear." Thus, when the context is felicitous and provides pragmatically proper expectancies for the use of negation, it is expected that listeners would process negative elements fast and incrementally in sentence interpretation.

Similar findings surrounding the incremental processing of negation have been found for other languages, including those languages where negation differs in linear placement in the sentence. In Korean, for example, an SOV language where negation can either appear directly before or after the verb, Nam (2016) observed evidence consistent with speakers representing the positive counterpart when no felicitous context has been provided, while Lee (2017) found that pragmatic felicity could help negation processing by establishing expectancies for using negation.

So, if most negation use and processing occurs in contexts where listeners anticipate its use, then there is perhaps less reason for speakers to avoid delaying negation on their listeners behalf. If audience design is a factor in motivating the Neg-First principle, then it would be to avoid the 'worst case scenario' in negation processing.

It has also been proposed that prohibitive sentence types may be even more affected by functional considerations. Horn (1989: 449-450) argues that the Neg-First principle is likely to

be even more pertinent in prohibitive sentences than declaratives, since uncertainty about the polarity of a postverbally negated directive could have dire consequences. This argument echoes Jespersen's (1917: 5–6) that, in the case of prohibitions, "it is important to make the hearer realize as soon as possible that it is not a permission that is imparted". For an illustration in the extreme, Van Olmen (2010) notes that an impatient addressee of (20a) may do precisely the opposite of what the speaker wants, if they anticipate the positive directive in (20b) instead.

(20) Dutch (Van Olmen 2010: 486)

a. *Vermoor haar nou* ***niet****!*
kill.IMP her PRTCL[NOW] NEG
'Don't kill her!'

b. *Vermoord haar nou toch!*
kill.IMP her PRTCL PRTCL[still]
'Just kill her already!'

In order to investigate whether the impact of Neg-First is in fact greater in prohibitions over declarative sentences, Van Olmen (2010) investigated a geographically and genealogically balanced sample of 179 languages. In most of the languages, the negator in standard declaratives and in the prohibitive occupied the same position with respect to the main verb, 96 languages had preverbal negation in both clause types and 48 had postverbal negation in both clause types.

Languages in which the position of negation relative to the main verb differ between the two clause types might lend more insight into the hypothesis that the Neg-First principle is stronger in prohibitions than in declaratives, however Van Olmen (2010) did not find strong evidence for this. Rather, he found that the number of languages in the sample with preverbal negation in the declarative and postverbal negation in the prohibitive (11 languages) is nearly the same as the number of languages with the opposite distribution (12 languages). Nadëb (example (21)) demonstrates preverbal negation in both its basic clausal negation strategies but postverbal negation in the prohibitive, whereas Iraqw (example (22)) shows the reverse pattern. Note that

the negative marker in the prohibitive is often a special marker dedicated to that construction (van der Auwera 2006), and these are glossed as prohibitive markers.

(21) Nadëb (Weir 1994: 295, 296, 305)
a. ***Dooh*** *kalapéé a-ód.*
NEG child PFX.cry.NIND
'The child isn't crying.'
b. ***Na****-ód* *kalapéé.*
NEG.cry.NIND child
'The child isn't crying.'
c. *Mi-~~ʉ~~g* ***manih****!*
2SG.ASP-drink.NIND PROH
'Don't drink (it)!'

(22) Iraqw (Mous 1992: 164, 168)
a. *Aníng a doohl-a-**ká**.*
I COP.1/2 cultivate-INF-NEG
'I don't cultivate.'
b. *M-a* *doohhl-**aar**!*
PROH.COP.1/2 cultivate-PROH
'Don't cultivate!'

While his findings do not seem to support the hypothesis that Neg-First is even stronger in prohibitives, Van Olmen (2010) notes that there were twelve languages in the sample that could not be readily categorized because of lack of information on actual usage, and among these twelve, nine may show signs in favor of special prominence of the Neg-First principle in prohibitive speech acts. For example, Nivkh marks basic negation postverbally, either with a negative auxiliary ḵ‘au or with the incorporated negative verb kavr/ġavr, as shown in (23a and 23b). The prohibitive can be formed with this latter negative verb, as in (23c), or with the preverbal prohibitive marker t‘a, as in (23d). However, it is not clear how these forms differ in use or which form is the most regular, so these were not categorized.

(23) Nivkh (Gruzdeva 2001: 62, 68)
a. *If p‘ry-doχ¸**k‘au**-d’.*
he come-DAT NEGAUX-IND
'He didn't come.'

b. *Jaŋ p'ř-**ǧavr**-d.*
he come-NEG-IND
'He didn't come.'
c. *Ra-**ǧavr**-ja!*
drink-NEG-IMP.SG
'Don't drink!'
d. ***T'a** ra-ja!*
PROH drink-IMP.SG
'Don't drink!'

Another form of evidence would be relevant if Neg-First is understood to be a general preference to have the negator be 'as early as possible' instead of merely before the verb. This has not been systematically checked yet to my knowledge, but in this case, a pattern such as that found in Cabo Verdean Creole could also indicate a preference for Neg-First principle in prohibitions above and beyond declaratives. Example (24) demonstrates the Cabo Verdean Creole pattern in which the negator in both declarative and prohibitive constructions occurs before the verb, but in prohibitive sentences, the negator must precede the subject clitic instead of following it as in declarative sentences. The prevalence of this type of pattern would have to be compared against that of the inverse, where the negator is placed later in the prohibitives than declaratives, even if it were on the same side of the verb in both clause types.

(24) Cabo Verdean Creole (Baptista 2002: 260)
a. *Bu **ka** odja algen.*
2SG NEG see someone
'You did not see anyone.'
b. ***Ka** bu bai!*
NEG 2SG leave
'Don't leave!'

Overall, the claim that the Neg-First principle plays a particularly strong role in the prohibitive relative to the declarative could have some validity, but ultimately, factors related to basic word order and/or the origin of prohibitive markers (negative markers associated

specifically with the prohibitive) probably also play a large role. I will expand on the role of grammaticalization, in both standard negation and prohibitives further in the next section.

In summary, it has been proposed by several scholars that functional considerations may help explain the Neg-First Principle, because delaying negation may result in processing difficulty or miscommunication. However, the evidence for this hypothesis is somewhat weak or uncertain.

2.2.3 Grammaticalization Processes

The final factor which has been proposed to play a role in the order of negation, either in concert or in conflict with the Neg-First principle, is the role of the natural diachronic process of grammaticalization. One way in which negators may form is by the process commonly known as Jespersen's Cycle. In this process, an original negative marker is phonologically weakened and/or loses semantic force, eventually requiring the support of an additional negative marker which eventually subsumes the original.

Jespersen's Cycle has famously been observed in French, shown in example (25) which demonstrates the way that the expression 'I don't say' has changed over time from Latin to present-day French. The Classical Latin negator *non* became *ne* in Old French and was emphasized with a second element *pas*, originally meaning ‘footstep’. In modern, colloquial French the original negator ne is dropped entirely.

(25) Jespersen's Cycle in history of French (Zeijlstra 2016)
   a. Classical Latin: ***Non*** *dico*
   b. Old French: *Jeo* ***ne*** *di*
   c. Middle French: *Je* ***ne*** *dis* ***(pas)***
   d. Modern French: *Je* ***ne*** *dis* ***pas***
   e. Colloquial French: *Je* ***(ne)*** *dis* ***pas***

A similar progression is observed in English, as shown in example (26), which lays out the progression of the expression 'I don't say' over time in English. In 11th and 12th century Middle English the preverbal negator *ne* became obligatorily supported by the element not, derived from the Old English indefinite pronoun *nawiht* (no thing). By the 15th century *not* was fully responsible for the expression of negation, and do-support entered the language giving rise to the present day system in which *not* or *n't* follows the auxiliary verb and precedes the lexical verb in negative sentences.

(26) Jespersen's Cycle in history of English (Willis 2016)
a. Old English: *ic* ***ne*** *secge*
b. Middle English: *I* ***ne*** *saye* ***not***
c. Early Modern English: *I say* ***not***
d. Modern English: *I do****n't*** *say*

Jespersen's Cycle has been observed less frequently outside of the languages of Western Europe and North Africa. This may be because it is less common in these languages, but is also certainly influenced by the difficulty of establishing whether a language has undergone Jespersen's Cycle without a strong textual record to reference. However, Jespersen's Cycles have been proposed and identified for numerous non-Indo-European languages, including Niger–Congo languages (Güldemann 2011: 117), Bantu languages (Devos & van der Auwera 2013), and Athabaskan languages (van Gelderen 2008), speaking to the widespread regularity of this process. For example, Pak (2020) describes Jespersen Cycle effects in indicative main clauses of Logoori, a Bantu language in the Luyia subfamily spoken in western Kenya, summarized in example (27).

(27) Jespersen's Cycle in history of Logoori (Pak 2020)
a. Proto-Bantu: ***si****-VERB*
b. 20th Century Luyia: ***si****-VERB* ***daave***
c. Modern Loogori: VERB ***daave***

According to Pak (2020), Logoori inherited a negative verbal prefix *si-* from Proto-Bantu which was once the sole negator, and this negation style is observed in a 1951 Bible and in traditional song. At some point, a clause-final negator *daave* was adopted into the language, perhaps a borrowing of the negative interjection *dawe* from the neighboring language, Luo, and used in addition to *si-*. By 1947 it was described as part of the negative construction in Appleby's grammar of Luyia, and Kanyoro's 1983 grammar of Logoori describes all three variations in (27) as grammatical in Logoori. Pak reports that her speaker-consultant rarely used *si-* in indicative main clauses, instead using *daave* in both elicited and narrative speech, and this pattern is confirmed by other contemporary sources, indicating a further progression of Jespersen's Cycle in these clauses.

In all of the above examples, we see a preverbal negator becoming replaced by a postverbal negator, which originated as an emphatic element. Negators tend to develop from two types of emphatic elements used in negative contexts: *minimizers*, which denote very small amounts of something (e.g. *a sip* in I didn't drink a sip), and *generalizers*, which extend the domain of quantification (e.g. *anything* in I didn't eat anything). We would expect such elements to occur in object position—so in VO languages, they would appear after the verb, and in OV languages, they would appear before the verb. Therefore, the common route for grammaticalization of negation particles seems to work in competition with a Neg-First bias.

Dahl (1979) notes that in his typological sample, postverbal placement of uninflected negation particles seems to be restricted to two groups: Germanic and Romance languages, and West African languages mostly belonging to the Niger-Congo Family. Dahl (1979: 95) speculates:

> [Among Germanic and Romance languages,] it is noticeable that the postverbal placement of Neg is a relatively recent phenomenon, which has arisen through Jespersen's Cycle. [...] It

> is thus tempting to assume that preverbal Neg placement is a natural universal tendency which may be disturbed by an equally natural diachronic process, viz. Jespersen's Cycle. If this is the case, one would expect the languages where this has happened to exhibit tendencies to return to preverbal placement. Some evidence for this can be found in the modern Scandinavian languages.

The case of English, where the loss of V-to-T raising and emergence of do-support returned the negator to a position before the lexical verb is a further example of a postverbal negator becoming preverbal.

Unlike negative particles, negative auxiliary verbs are more likely to precede the lexical verb in VO languages and follow the lexical verb in VO languages. This tendency is particularly important for the development of prohibitive markers, which tend to grammaticalize from verbs meaning to stop or abstain. A prohibitive construction is said to be 'compositional' if it consists of the positive imperative form in conjunction with the standard negator that appears in negative declarative sentences. But there is a worldwide preference for *non*-compositional prohibitives; the prohibitive was found to be non-compositional in more than three quarters of the 496 languages in van der Auwera & Lejeune's (2013) sample, such that the main prohibitive construction is formed using a negator not found in standard negative declarative sentences, a verb form not found in standard imperative sentences, or both. Example (28) illustrates the Ju|'hoan prohibitive, formed with the positive imperative verb form alongside a dedicated negative marker.

(28) Ju|'hoan (Snyman 1970: 140–145)

a. *Tšhi !xeri!*
   drink liquor
   'Drink the liquor!'

b. *Mi* |***óá*** *!hún n!haì.*
   I NEG kill lion
   'I don't kill the lion.'

c. ***N***||***a*** *tšhi !xeri!*
   PROH drink liquor
   'Don't drink the liquor!'

These special prohibitive strategies commonly result from the grammaticalization of lexical items or phrases for abstaining or stopping into dedicated prohibitive negators (van der Auwera 2006, Heine & Kuteva 2002). In Korean, prohibition is expressed with the imperative form of the auxiliary verb *malta* meaning 'stop', following the connective suffix *-ci*, as in (29c). This prohibitive construction follows the main verb, as is typical for auxiliary verbs in SOV languages like Korean. Recall that Korean has both a postverbal (29a) and preverbal (29b) method for expressing negation in standard declaratives; the postverbal method involves the auxiliary verb *ha*.

(29) Korean (Chang 1996: 101, 103; Rhee 2003)

a. *Yong-i TV-lul po-ci **an**-ha-yo.*
Yong-SBJ TV-OBJ see-SUSP NEG-AUX-POL
'Yong doesn't watch TV'

b. *Yong-un TV-lul **an** pwa-yo.*
Yong-TOP TV-OBJ NEG see-POL
'Yong doesn't watch TV.'

c. *Ka-ci **mal-la**!*
go-SUSP stop-IMP
'Don't go.'

Korean therefore follows Dahl's (1979) prediction that in verb-final languages, negative auxiliaries should tend to follow the verb, like auxiliaries in general, rather than be subject to the preverbal tendency.

Overall, the grammaticalization pathways known to generate negators in language, help to explain other patterns of negation ordering across languages, but do not necessarily help to explain a general Neg-First preference, because they predict opposite outcomes relative to the verb for OV and VO languages. Emphatic particles will tend to be derived from elements in object position, whereas negators associated with or derived from auxiliary verbs will tend to appear on the opposite side of the verb from the object. The prevalence of preverbal particles in both OV and VO languages is therefore of interest and may be the result of a separate Neg-First

preference, perhaps motivated by the previously explored functional or processing considerations.

2.3 Summary

Dahl (1979: 96), discussing evidence from acquisition suggesting that preverbal negation may be an 'unmarked' order, wrote: "Of course, one does not explain a linguistic phenomenon just by labelling something as 'unmarked' relative to something else. At present, I can see nothing that would provide a real explanation of the preverbal tendency for negation". In the years since, several explanations have been proposed which could account for the preverbal tendency for negation and other patterns of negation order. These proposed explanations emphasize to a different extent the role of constraints and pressures affecting language acquisition, production, comprehension, and grammaticalization. In the following chapters, I present artificial language learning experiments in which participants learn a language with both pre- and post-verbal negation to investigate the presence of a bias towards preverbal negation in acquisition among learners with different language backgrounds, and whether interaction with an interlocutor increases use of preverbal negation.

Chapter 3 Connecting Typology and Individual Biases

As discussed in Chapter 2, evidence from naturalistic first and second language acquisition suggests that there are trajectories in early stages of learning in which learners systematically produce negation earlier in the sentence than it would occur in adult-like or target-like sentences, particularly when the position of negation is variable or opaque in the target language. However, there is still a great deal of uncertainty as to the cause of such trajectories and whether they can be explained by a unified preference for leftward placement of negation, or by separate factors specific to any given language under consideration. For example, tendencies which show greater use of the preverbal (short-form) negation construction in Korean over the alternative postverbal (long-form) negation construction (Kim & Yun 2013), could be driven by greater exposure to the preverbal construction in child- and learner-directed speech (Lee 2008, Kim 1991) or by a preference for particle negation over morphological negation, rather than a preference for preverbal negation per se. Because natural language data cannot control for variation in learners' linguistic input and structural differences between languages, scholars have utilized artificial language learning paradigms to detect and identify biases present in language learning and use.

3.1 Previous Artificial Language Learning Studies

Artificial language learning paradigms allow for researchers to have greater control over the learners' linguistic input and learning conditions, comparing patterns of interest while minimizing confounding factors. Studies using artificial language learning with both children and adults have revealed biases which can provide potential explanations for typological patterns

and trajectories of language change; artificial language learning studies can also allow us to explore how biases affecting language learning change or appear across development, how they may be amplified or diminished by language experience, and how they may apply to other cognitive domains. For example, in the domain of syntax and word order, Culbertson et al. (2012), find evidence of a learning bias parallel to Greenberg's (1963) Universal 18—Adjective-Noun order implies Numeral-Noun order. Indeed, languages with Adjective-Noun and Noun-Numeral order are exceedingly rare (4% of languages in WALS).

To date, only a few artificial language learning studies have incorporated or investigated standard negation, and none has examined preferences regarding the ordering of the negator directly. Baptista et al. (2014) incorporated standard negation, training learners on a miniature language that contained morphemes for negation and pluralization to investigate the nature of convergence and transfer in second language acquisition. They found that English-speaking participants produced sentences in the language most accurately when the negative and plural morphemes had both similar phonological form and semantic function to the corresponding English morphemes. Of note, one type of error made by participants of this study regarded the placement of the negative morpheme; participants in all conditions occasionally produced the negative morpheme preverbally, even though the negative morphemes in the artificial language used in the study were exclusively postverbal. In another study, Maldonado & Culbertson (2021) used an artificial language learning paradigm to examine whether learners are sensitive to a correlation between the morphological status of the negative marker (affixal vs. adverbial) and the interpretation of the sentence with two negative elements (negative concord vs. double negative). Their findings failed to find evidence of such a connection, but instead suggest that learners find it easier to learn negative concord languages compared to double negation

languages regardless of whether the negative marker is an adverb or affix, in line with evidence from natural language acquisition.

Artificial language learning paradigms offer the unique opportunity to narrow in on specific sources of learning and production asymmetry. Though artificial languages used for experimental studies lack the rich structure, context, and complexity of natural languages by design, previous research demonstrates that there are important similarities in how learners process artificial and natural languages (Amato & MacDonald 2010, Fehér et al. 2016). Converging evidence using a variety of methods and sources is critical to understanding the mechanisms at play in natural language learning and further, the mechanisms which link biases of individuals to patterns of language in the larger community. In Chapter 2, I outlined evidence from typology, language emergence, and language acquisition, and presented possible mechanisms affecting acquisition, production, communication, and grammaticalization. In this chapter, and the rest of this dissertation, I take initial steps towards testing the presence of Neg-First biases in the early stages of second language acquisition by examining whether participants exposed to an artificial language containing both preverbal and postverbal negation demonstrate a bias to overproduce preverbal negation compared to their training input.

The experiments in this dissertation utilize an artificial language learning paradigm developed by Hudson & Kam (2005, 2009) and Culbertson et al. (2012) which involves training learners on a language that allows for free variation between two or more alternative structures. For example, more than one word might map to the same meaning, or more than one word order might be used to describe the same scene. This paradigm, referred to by Culbertson (2021) as the 'regularization' paradigm, does not ask learners to innovate new structures they have not already been exposed to, but idealizes the input that a learner might be exposed to during a period of

linguistic contact or change when several variants are present in a language, in free variation, or constrained probabilistically by social or linguistic factors that may be opaque to a beginner language learner. Evidence for a bias comes from observing the proportion with which learners use a given variant compared to its frequency in the input; using one of the patterns more frequently (i.e., regularizing or overproducing a pattern) is taken to indicate a preference for that pattern.

3.2 Experiment 1: Investigating Neg-First Biases Among English Speakers

Experiment 1 aimed to determine whether English-speaking language learners demonstrate evidence for a bias towards preverbal negation in the early stages of learning a new language.

3.2.1 Methods

Participants. Adult participants who spoke English as a first language (operationalized as learned before age three) were recruited from the Ann Arbor area through classrooms and flyers on and around the university campus. Data from 60 participants were collected and analyzed (three additional participants took the study but were excluded from analysis for failing to meet an accuracy threshold of 75% in the final production phase).

Materials. The artificial language includes six nouns, four verbs, and a negation particle (see Table 5) which are combined to create simple transitive sentences. The language used SVO word order and no noun was used in the same sentence as both a subject and an object, so a total of 120 total affirmative sentences were created from the lexicon. The negative particle could appear either directly before or after the verb. The proportion of preverbal negation ordering was manipulated depending on the participant's training condition.

Table 5. Artificial language lexicon for Experiment 1 with broad IPA transcriptions.

Nouns			**Verbs**		
boxer	daki	/dɑki/	shoot	patu	/pɑtu/
chef	falit	/fɑlɪt/	kick	sal	/sæl/
burglar	ludin	/ludin/	point	toma	/toʊmə/
cowboy	mook	/muk/	punch	umi	/umi/
doctor	pepo	/pipoʊ/	**Negation Particle**		
artist	sido	/sidoʊ/	'not'	pik	/pɪk/

Audio of these sentences were synthetically generated using Amazon Polly's 'Salli' voice.[6] Each sentence was paired with a cartoon image (see Figure 2).[7]

Figure 2. Example cartoon stimuli for Experiment 1.

Design & Predictions. Participants were randomly assigned to one of three input conditions. For each condition, the training data differed only in the proportion of negative sentences which placed the negator preverbally rather than postverbally. Participants in the Majority NegV Condition were exposed to majority (75%) preverbal negation ordering, participants in the Majority VNeg Condition were exposed to majority (75%) postverbal negation ordering, and participants in the Equiprobable Condition were exposed to an equal proportion of both word orders.

[6] https://aws.amazon.com/polly/
[7] Cartoon images were by Sara Rolando and provided by Kenny Smith

The dependent measure of interest is the proportion of use of the majority pattern of preverbal negation in productions of participants across the different experimental conditions. The bias to regularize unconditioned variability in the input (Hudson Kam & Newport 2005, 2009; Culbertson et al. 2012) may push learners to use the majority pattern more frequently than it is found in the input. A bias in favor of preverbal negation predicts that participants in the majority preverbal negation condition will boost the proportion of preverbal negation in their productions relative to the input to a greater degree than participants in the postverbal negation condition. We also might expect that if there is a strong bias towards preverbal negation, any frequency in the equiprobable condition would be in the direction of preverbal negation, if frequency boosting takes place at all.

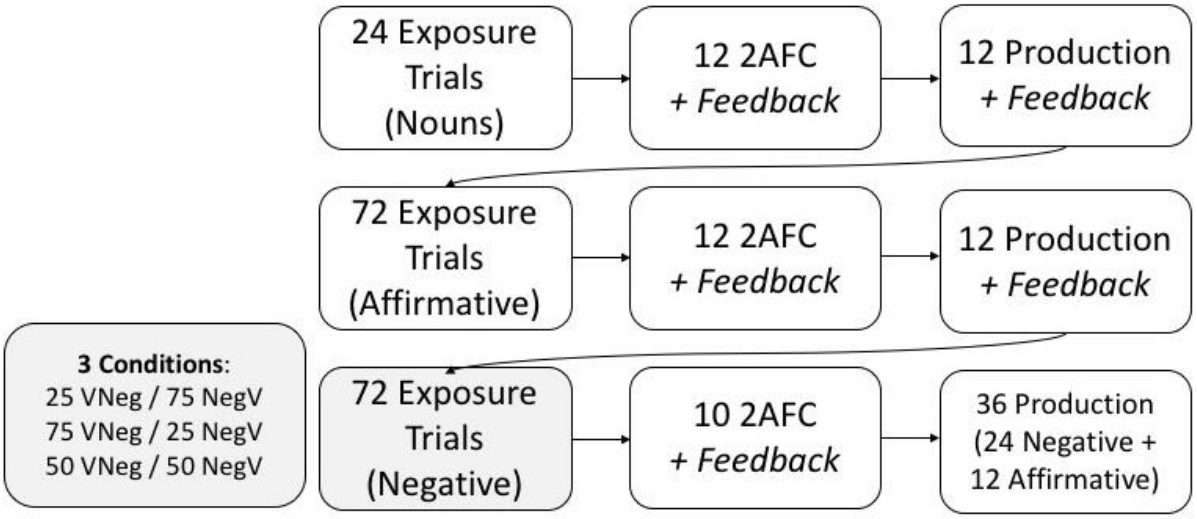
Procedure. Each participant was trained and tested on the artificial language in a single 60-minute session. The experiment was implemented in PsychoPy (Pierce et al. 2019) on a computer in a private room, and participants wore headphones through which audio was played. The experimenter remained in the room with the participant for the duration of the experiment. The experiment began by introducing participants to the learning scenario. Participants were told that they would be learning a new language, that this language may not always form sentences the way that English does, and it was their task to learn how to form sentences in this language. Participants were also informed that they may not take notes during the experiment. The experiment consisted of three distinct exposure phases, each followed by a brief testing phase.

Participants were trained on the artificial language in three phases, which introduced nouns, affirmative sentences, and negative sentences respectively. An overview of the design is presented in Figure 3. In exposure phases, participants viewed a cartoon image with the corresponding text in the artificial language printed beneath it and heard the corresponding audio

through a pair of headphones. In two-alternative forced choice (2AFC) trials, two images were presented on the left and right side of the screen, above text in the artificial language.

Participants were told to use the left and right arrow keys to choose the image on the screen that corresponded to the text. In production trials, participants were presented with an image in the center of the screen and asked to type the text that described the image into a text box. A word bank with the relevant vocabulary presented so far was provided in the upper left corner of the screen for participants to reference.

Figure 3. Experiment 1 procedure overview



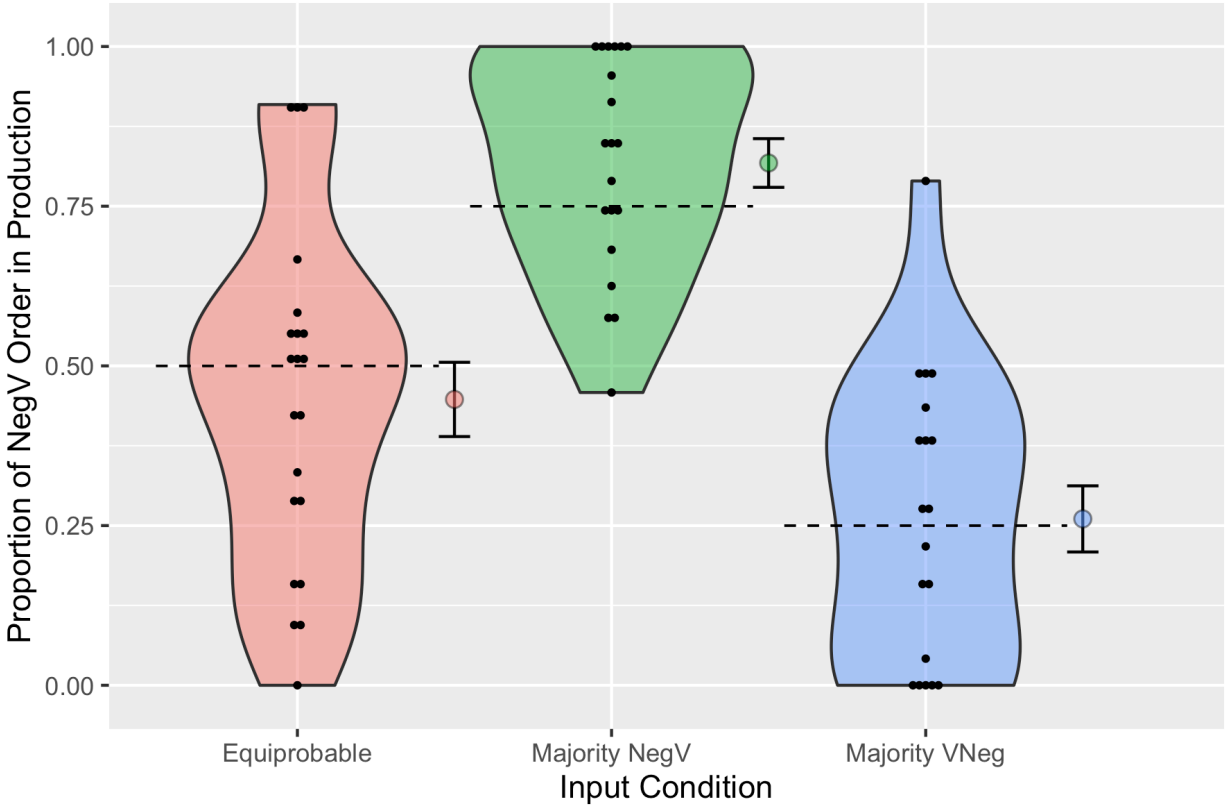
3.2.2 Results

The data files and analysis code for this experiment are available at https://github.com/DaniLBurgess/BurgessDissCh3. Statistical analyses were conducted in R, and mixed effects logistic regression models were fit using the lme4 package (R Core Team 2021, Bates et al. 2015).

The average proportion of productions with preverbal negation by condition are shown in Figure 4. This figure and all analyses in this section include only accurate production trials where

the correct vocabulary items were produced in an order that was modeled during training. Roughly seven percent of trials (107 out of 1440) were excluded due to inaccuracy. Accuracy in forced choice task trials were high (>95%) across phases and conditions.

Figure 4. Proportion production of preverbal negation order by condition in Experiment 1. Dashed lines indicate proportion of preverbal negation order in the input; small points show subject means; the large points offset to the right of each violin plot indicate the overall mean for each condition.



To examine whether the proportion of preverbal negation productions in the Equiprobable Condition was significantly different than the input proportion of 50%, I ran a model with Preverbal Negation word order use as the binary dependent variable, Condition as a fixed effect (reference level = Equiprobable), and random intercepts for Subject.[8] This model produced a non-significant intercept ($\beta = -0.29$, $z = -0.89$, $p = 0.38$), indicating that the

[8] A model that also included random intercepts for Item was over-fitted.

probability of productions with NegV word order was not significantly different from chance in this condition. In order to test the prediction that participants in the Majority NegV Condition boosted the majority order significantly more than those in the Majority VNeg Condition, another model was run with Majority Order word order use as the binary dependent variable, Condition as a fixed effect (reference level = Majority NegV), and random intercepts for Subject and Item. This model revealed that the difference in the use of majority word order between the Majority NegV and Majority VNeg conditions was not significant ($\beta = -0.52$, $z = -1.03$, $p = 0.31$).

3.2.3 Discussion

The results of Experiment 1 are inconclusive with regards to demonstrating the existence of a learning bias that favors preverbal negation. Learners in the majority preverbal negation condition did boost the majority word order in their productions more than the learners in the majority postverbal negation condition, as would be predicted by a learning asymmetry favoring preverbal negation, but this difference in boosting was not statistically significant. Furthermore, the participants in the Equiprobable Condition, did not produce a significantly different proportion of preverbal negation overall than they received in the training input.

There are various ways that one can interpret this result. It could be that there truly is no learning asymmetry favoring preverbal negation in adult English-speakers. In this case, it may be that the observed typological tendency towards early negation came about by chance or by grammaticalization processes that do not relate to any independent advantage for early negation. Or, there may yet be an asymmetry influencing language change and typology that appears not in learning (perhaps as a result of constraints on memory encoding and retrieval as learners first become familiar with a system), but in the process of cultural transmission and communication.

Consider Dryer's (1988) hypothesis that a Neg-First principle is grounded in the communicative desire to reduce the potential for the listener to misunderstand or be led down a 'semantic garden path' by delaying negation until later in the sentence. The current study, which did not elicit linguistic production in an interactive, communicative context, fails to truly probe this motivation. Experiment 2, reported in Chapter 4, aims to investigate whether negation-ordering asymmetries appear in communication (above and beyond learning and production in isolation).

Because English has a system in which negation always appears after the finite auxiliary verb but before a lexical verb, it is not straightforward to predict how a preference for English-like ordering may bias learning of a miniature language like the one used in this study which included no auxiliary verbs. It could be the case that a bias for post-finite negation on the basis of English negation order opposes a general bias for negation to precede the verb for some learners in this study. It is also the case that while typological databases tend to report the position of negation relative to the *lexical* verb, Horn (1989) speculates that the preverbal negation tendency is the tendency for negative markers to precede the *finite* verb. Experiment 3, reported in Chapter 5, aims to investigate the extent to which negation-ordering asymmetries are mediated by participants' language background by comparing the English and Japanese-speaking learners of similar artificial languages.

3.3 Summary

Experiment 1, though inconclusive as to the presence of a bias favoring preverbal negation among English speakers in the early stages of learning a new language, was an important first step into exploring whether and under what circumstances preverbal negation is treated differently from postverbal negation, and how this could (at least partially) explain the typological tendency for preverbal negation. Acknowledging that there was a small non-

significant trend towards learners boosting the majority word order in the preverbal negation condition over the postverbal negation condition, another interpretation of the results is that there is a true learning bias underlying this behavior, but that this bias is sufficiently weak that this study was underpowered to detect it, or the task was not sufficiently difficult to encourage this tendency.

Because artificial language learning is still a relatively new methodology with a rapid development and a large amount of variation in design and procedure across experiments, it is still not very well understood how methodological differences such as the size of the artificial language, the length of the exposure period, the nature of exposure and testing procedures, and participant demographics and language backgrounds affect how cognitive biases play out in artificial language learning outcomes. However, Hudson Kam & Chang (2009) found that increased demands on lexical retrieval led to increased regularization in adult learners. This is consistent with more recent research indicating that the bulk of regularization behavior may be associated with limitations associated with production, rather than with encoding frequency in learning (Ferdinand et al. 2019). Experiment 1 included a word bank during the production task to make the experiment more tractable for participants to complete in a single training session, however the inclusion of the word bank and the lack of time pressure in the production phase of the experiment may have made the task too easy to induce regularization behavior in participants. Due to the concern that the inclusion of the word bank limited the amount that participants would regularize, obscuring a bias that might emerge only when there is increased difficulty or demand on lexical retrieval, this memory aid was removed from Experiments 2 and 3.

Chapter 4 What Role Does a Communicative Context Play in Neg-First Biases?

In this chapter, I review the hypothesis that audience design, the practice of shaping language to accommodate the addressee, plays a role in motivating a Neg-First bias in communication, which in turn could shape language typology and the cross-linguistic prevalence of preverbal negation. I then present Experiment 2, which combines an artificial language learning paradigm with a dyadic director-matcher paradigm in order to test this hypothesis; the inclusion of a dyadic interaction task enables us to determine whether a preference for early negation is more prevalent in a context in which information is being transmitted to a communicative partner, compared to production in isolation.[9]

4.1 Audience Design Proposal For Neg-First

As discussed in §2.1.1, several typological surveys have demonstrated that the standard negator is usually adjacent to the finite verb and tends to precede the lexical verb across the world's languages. This preverbal tendency for sentential negation marking is relatively independent of main constituent order, although it appears that the tendency is strongest for verb-initial languages and weakest for verb final languages. Dryer (1988) proposed that a combination of three sometimes-competing principles may be able to account for the broad patterns of negation and constituent order found in a geographically and genetically balanced sample of 345 languages: the Branching Direction Principle (which states that languages have a tendency to order heads and phrases consistently, e.g., verb-final languages tend to also use postpositions),

[9] This chapter is adapted from a paper published proceedings paper (Burgess 2022), though the results and discussion have been greatly expanded here.

the Negative-Plus-VO Principle (which states that the negator should not interrupt the VO unit), and the Negative-Before-Verb (i.e., Neg-First) Principle. In combination, these three principles could motivate why the tendency to use preverbal negation was stronger in VO languages than in OV languages.

The three principles proposed by Dryer (1988) were not meant to represent inviolable constraints on languages, but general preferences motivated, perhaps, by advantages in processing. For example, correlations in ordering for different types of phrases (e.g., the tendency for OV languages to have postpositions), which Dryer (1988) groups under the Branching Direction Principle, have been argued to allow for efficient online processing (e.g., Hawkins, 2004). Dryer (1988) also speculated about a potential functional motivation for a Neg-First principle in that "delaying [negative morphemes] increases the risk of misunderstanding, [...] since the apparent meaning of a sentence up to but not including the negative will be the opposite of the intended meaning" (102). Horn (1989:449-450) took up a similar focus on audience-design in the role of the Neg-First principle, arguing that Neg-First was likely to be even more pertinent in prohibitive sentences than declaratives, as an impatient addressee may do the opposite of what the speaker wants if negation is delayed (though c.f. Van Olmen 2021, who found no significant difference in the prevalence of preverbal negation between declarative and imperatives).

Potentially supporting the audience design proposal, psycholinguistic research has demonstrated that positive counterparts to negative sentences are often represented in the early stage of processing, particularly when contextual support is absent (see §2.2.2., Tian & Breheny, 2019 for a review of the negation processing literature). However, when the context is felicitous and pragmatically proper expectancies for the use of negation are provided, there is less of a

tendency to first consider the positive argument. Therefore, if audience design is a factor motivating the Neg-First principle, then it would be to avoid a 'worst case scenario' in negation processing, when contextual cues to negation are lacking or not well-perceived.

Though there has been extensive research concerned with the processing costs of negative sentences in comparison with affirmative sentences, the specific effect on processing of delaying cues to negation until after the verb has not been widely investigated. However, Lipski (2018) did report that among Spanish-Palenquero bilinguals, languages which crucially differ in the placement of the sentential negator (preverbal in Spanish, clause-final in Palenquero), processing of Palenquero clause-final negation was more vulnerable to the effects of cognitive distractions. This provides evidence that delayed negation may make online processing more precarious in less-than-ideal conditions (e.g., in the presence of high noise, distraction, cognitive load).

4.2 Artificial Language Learning & Dyadic Interaction

Previous research has combined ALL paradigms with communicative tasks (typically a dyadic director-matcher paradigm, in which participants trade off describing and selecting images) to investigate the role of communication in regularization behavior. Fehér et al. (2016) report an experiment in which participants were trained on a language with two equiprobable word orders, both of which were verb initial but differed in the order in which the agent and patient of the verb were expressed. After training, participants then performed a director-matcher task either in a dyad (with another human participant), in a pseudodyad (with a computer they believed to be another human participant), or solo (with a computer with no deception). Fehér et al. (2016) found evidence of structural priming regardless of whether their partner was a human or computer: participants were more likely to repeat the immediately preceding word order

choice of their partner than use a different word order. However, structural priming only led to convergence (similarity in proportions of word order use) in human-human dyads, in which some dyads converged on highly regular systems of agent-first order. When paired with the computer, which continued to produce each word order with equal probability, participants in pseudodyads continued to increase regularity (towards more agent-first word order use) over the course of the interaction despite priming effects being present. Therefore, priming alone does not inevitably lead to convergence; reciprocal priming, in which participants prime each other in turn, may thus be an important factor leading to regularization and convergence.

Communication with a human partner also does not determine greater regularization alone. Saldana, Smith, Kirby, & Culbertson (2021) found that participants learning a different artificial language regularized unconditioned variation in the input to similar degrees when they produced phrases in isolation and during communicative interaction, suggesting that regularization is not fully driven by communication, at least in the context of L2 production. There are several differences between the studies of Fehér et al. (2016) and Saldana et al. (2021) that could explain these different results. The input languages in the study by Saldana et al. (2021) had many more variants and the baseline level of regularization by participants in isolation was already relatively high even before the communicative interaction phase. Furthermore, Saldana et al. (2021) observed no strong bias towards a specific pattern from the input, whereas Fehér et al. (2016) observed a bias towards the agent-first word order pattern. Though the participants in the Fehér et al. (2016) study showed an initial bias which was then reinforced during communication, the dyads in the Saldana et al. (2021) study were not closely aligned, regularizing the input in different ways, so the mixture of their production systems resulted in higher variability.

Finally, Fehér et al. (2019) found evidence of *grammar-based asymmetric accommodation* in an experiment using artificial language learning and interaction paradigms: variable users of optional linguistic variants were more likely to accommodate to categorical partners (i.e., users who always or never used an optional singular marker), but categorical users did not tend to accommodate variable partners by becoming more variable. Such asymmetric accommodation offers a mechanistic explanation for obligatorification during language change (i.e., the tendency for constituents to shift from occurring variably and being pragmatically conditioned to being obligatory and grammatically conditioned).

These studies demonstrate that the amount of regularization of linguistic variation (i.e. movement towards a more categorical system) is the result of a complex interaction of not only the nature of the production task, but the language structures in question, the number and proportion of variants, and whether there are any variant-specific biases.

Experiment 2 contributes to this body of literature exploring how biases in language learning and use interact to shape variation among natural languages, providing empirical evidence in the domain of the word order of standard negation.

4.3 Experiment 2: Investigating Neg-First Tendencies in Learning and Interaction

Experiment 2 investigates the presence of a preference for earlier negation among English speakers learning an artificial language that freely allows both preverbal and postverbal negation. This study also incorporates a dyadic interaction (director-matcher) paradigm, to investigate whether a preference for early negation could be plausibly driven by the specific pressures of communication, as opposed to learning and production in isolation. Besides this added dyadic paradigm, Experiment 2 diverges from Experiment 1 in several other respects: the word order of the artificial language is VSO instead of SVO to mitigate direct comparison to

English; the production stages do not include a word bank, in order to increase demands on lexical retrieval and thereby encourage adult learners to regularize more (Hudson Kam & Chang 2009); and, due to COVID-19 restrictions, participants completed the study remotely rather than coming in to the lab.

4.3.1 Methods

Participants. 54 adults completed the experiment, 18 participants in each of three randomly assigned input conditions forming 9 dyads per condition. 8 more participants were recruited but failed to finish due to technical difficulties (n=4) or failing to match with a partner within the time limit (n=4). All participants grew up speaking English at home and/or in school, lived in the US at the time of taking part in the study, and completed the study remotely on a computer. Participants were recruited through a combination of social media forums, flyers, and snowball sampling. Participants were paid $15/hour for their participation.

Materials. The artificial language lexicon and images were the same as that in Experiment 1. The included six nouns, four verbs, and negation particle which are combined to create simple transitive sentences (refer to Table 5). Audio of these sentences were synthetically generated using Amazon Polly's 'Salli' voice and each sentence was paired with a cartoon image depiction The language used VSO word order and the negative particle could appear either immediately before or after the verb. Whether a sentence appeared with pre- or postverbal negation ordering was not conditioned on any aspect of the sentence itself, but the proportion of each negation order in the input language varied depending on the input condition the participant was assigned to.

Unlike Experiment 1, VSO word order was chosen for Experiment 2 in order to mitigate transfer effects from experience with English. The word order was also chosen to avoid the

negator being placed adjacent to the object noun, so as to prevent participants from interpreting the variation in negation placement as a difference in the scope of negation (i.e., sentential negation vs. constituent negation).

Design & Predictions. Participants were randomly assigned to one of three language input conditions in which the proportion of preverbal and postverbal negation was manipulated: participants in the Majority NegV Condition were exposed to majority (75%) preverbal negation ordering, participants in the Majority VNeg Condition were exposed to majority (75%) postverbal negation ordering, and participants in the Equiprobable Condition were exposed to 50% of each ordering. During the testing phase of the experiment, each participant was paired with another participant in the same condition, creating 9 dyads in each condition.

The dependent measure of interest is the proportion of use of the majority pattern in productions of participants across the different experimental conditions. The bias to regularize unconditioned variability in the input (Hudson Kam & Newport, 2005, 2009; Culbertson et al. 2012) may push learners to use the majority pattern more frequently than it is found in the input. A bias in favor of preverbal negation predicts that participants in the Majority NegV Condition will regularize towards NegV word order to a greater extent than participants in the Majority VNeg Condition will regularize towards VNeg word order. We also might expect that if there is a strong bias towards preverbal negation, regularization in the Equiprobable Condition would be in the direction of preverbal negation.

The amount of regularization that occurs in the isolated production tasks and in the interactive task will also be compared. If the communicative context introduces a special preference for early negation, we predict participants to produce a greater proportion of NegV word order in the interactive condition than in the first production stage.

Procedure. Each participant was trained and tested on the artificial language in a single session no more than 90 minutes long. The study was conducted remotely with a custom program developed for the LIONESS experiment platform (Giamettei, Yehosseini, Gächter, & Molleman, 2020). Participants were run in even-numbered groups and given instructions over Zoom before being sent the LIONESS study link in Zoom. Participants were told that they would be learning a new language, that this language may not always form sentences the way that English does, and it was their task to learn how to form sentences in this language. They were also told that they would later be asked to use this language to play a game with another participant. Participants worked through series of three training stages in which they were exposed to the artificial language. This was followed by three critical production stages *(*recall 1, interaction, and recall 2), in which participants were asked to describe novel images withheld during the training phase in the artificial language.

Stage 1, Noun Training. Participants were introduced to the six nouns of the artificial language over 18 exposure trials. Then, participants were tested on their comprehension and production of these nouns with 12 trials of a two-alternative forced choice task (2AFC) followed by 12 production trials.

During this phase, and all training blocks to follow, the trials were presented as follows. On each exposure trial, participants were presented with an image and a corresponding description in the language, visually printed beneath the image and aurally presented. On each 2AFC trial, two images were presented on the screen above a description, and participants were directed to use the left and right arrow keys to choose the image that corresponded to the description. The correct answer was on the left side half of the time. Participants received feedback on their responses. For production trials, an image was presented in the center of the

screen and participants were asked to type the text that described the image into a text box. After participants submitted their answer, the correct answer appeared on the screen for 3 seconds regardless of participant accuracy.

Stage 2, Sentence Training. Participants were introduced to the language's four verbs in the context of complete transitive affirmative sentence over 48 exposure trials. Sentences were balanced so that each noun was presented equally often with each verb in both subject and object position. This was followed by 12 2AFC trials and 12 production trials not yet encountered by participants during the exposure stage. The distractor images in the 2AFC trials differed from the correct image by a single element (agent, action, or patient).

Stage 3, Negative Sentence Training. Next, participants were exposed to 48 negative sentences in the language. The images used in this phase resembles those images used for affirmative sentences, with a slightly transparent red circle with a diagonal line superimposed over the center of the image. As mentioned above, the proportion of preverbal vs. postverbal negation varied between experimental conditions. However, in each condition, sentences were controlled such that each negation order appeared with all verbs and all nouns in both subject and object positions, so that there were no unintended regularities that might appear to condition the order of negation in the sentence.

Participants completed 12 2AFC to test comprehension of negative sentences. For these trials, one of the images depicted the negated event, making this the *incorrect choice*, and the other image differed in the depiction of the patient (as in the image on the right side of Figure 5). Participants received feedback on this task as in previous tasks. If they chose the incorrect image, they received a reminder that "negative sentences tell you what did not happen". By having participants choose between 'affirmative' event drawings to test negative sentence

comprehension, we can check that participants understood the function of the negation marker without relying on the 'negation symbol' cue in the drawing.

Figure 5. Director-Matcher stimuli for Experiment 2. An example of a director input screen (left) and matching (2AFC) screen (right). For the sentence corresponding to, "cowboy not punch boxer," the correct match is the image in which the cowboy is punching the artist.

Stage 4, Recall 1 The production phase following negative sentence exposure and comprehension trials comprised 36 totals trials using images withheld from previous stages of the experiment (24 negative sentences and 12 affirmative sentences). Feedback was not provided to participants during this stage.

Stage 5, Interaction Production Participants were placed in a virtual waiting room and paired sequentially into dyads as they finished training in order to complete a director-matcher task. If a participant was not paired with a partner after 15 minutes, they were removed from the waiting room queue and paid for their time. In this task, participants alternated describing images for their partner, and selecting images based on their partner's description. One participant (the director) was presented with an image depicting a negated event and prompted to describe this image for their partner. In order to prevent communication in English or confusion caused by typos, the legal sentence (i.e., the sentence with the correct vocabulary items) that was the closest Levenshtein distance to the participant's input was transmitted to the matcher. Levenshtein distance is the minimal number of insertions, deletions, and symbol substitutions required to transform a string to another. The matcher had to choose the image that matched the description

from a display with two images. As in Stage 3, when the sentence was negative, one picture depicted the negated event while the other picture depicted a different patient than that in the negated sentence (see Figure 5 for an example of a director-matcher trial). After each trial both participants receive feedback (success or failure) and an updated score (‘‘Score so far: X out of Y”). Participants each described a pre-selected subset of 24 images (16 negative, 8 affirmative) chosen from the 36 images in the pre-interaction production stage, for a total of 48 trials in this stage**.**

Stage 6, Recall 2 Finally, participants viewed the same 24 images from Stage 5 (order randomized) and were asked to enter the appropriate sentence. No feedback was provided on the participants' responses.

4.3.2 Results

All statistical analyses were conducted in R, and mixed effects logistic regression models were fit using the lme4 package (R Core Team 2021, Bates et al. 2015). The data files and analysis code for this experiment and Experiment 5 are available at https://github.com/DaniLBurgess/BurgessDissCh4Ch5.

Prior to production response analysis, typos were automatically corrected by splitting each response typed by a participant into a series of words by splitting the string at the spaces and identifying the closest legal word for each. Levenshtein string edit distance was used to determine the closest match. If a participant's word was one Levenshtein edit distance away from only one legal word, then the participant's response was corrected to that legal word. Then, the word order for negative sentence trials (NegV vs. VNeg) was automatically labeled based on the closest match when comparing the whole corrected response to each of the correct target sentences. Responses which were greater than 5 edit distance away from a correct response was

labeled as inaccurate and excluded from analysis; responses where there was not a clear closest match to either target sentence were also excluded from analysis. The criterion of 5 edit distance was chosen because the maximum word length in the artificial language is 5 characters; where one word was incorrect or swapped for another, word order choice could still potentially be extracted.[10] Recall that during the interaction stage, any typos or errors given by the matcher were automatically corrected to the closest possible legal response when displayed to the matcher, so incorrect labels were never transmitted to the partner.

Accuracy. Accuracy in both the forced choice task and production tasks were high overall: 96.4% across forced choice task trials and 97.6% across production trials. All 54 participants that completed the experiment met a threshold of 75% accuracy in the final production (Recall 2) phase. Across critical negation trials in all three phases, the mean accuracy was 98.7%; 42 out of 3120 total trials across participants were excluded from further analysis. Figures showing accuracy information from Experiment 2 across different phases and conditions are available in Appendix C.

To check whether there were differences in accuracy in either the forced choice or production tasks in different input frequencies condition, the buildmer package in R was used to evaluate the maximal feasible model for each of forced choice task accuracy and production accuracy, and to subsequently identify the model with the best fit through automated backwards stepwise elimination (Voeten 2022). This process compares the fit of a more complex model to the fit of a nested simpler model, using the Likelihood Ratio Test to determine whether the effect of a covariate is significant. The maximal effect structure for both the forced choice task

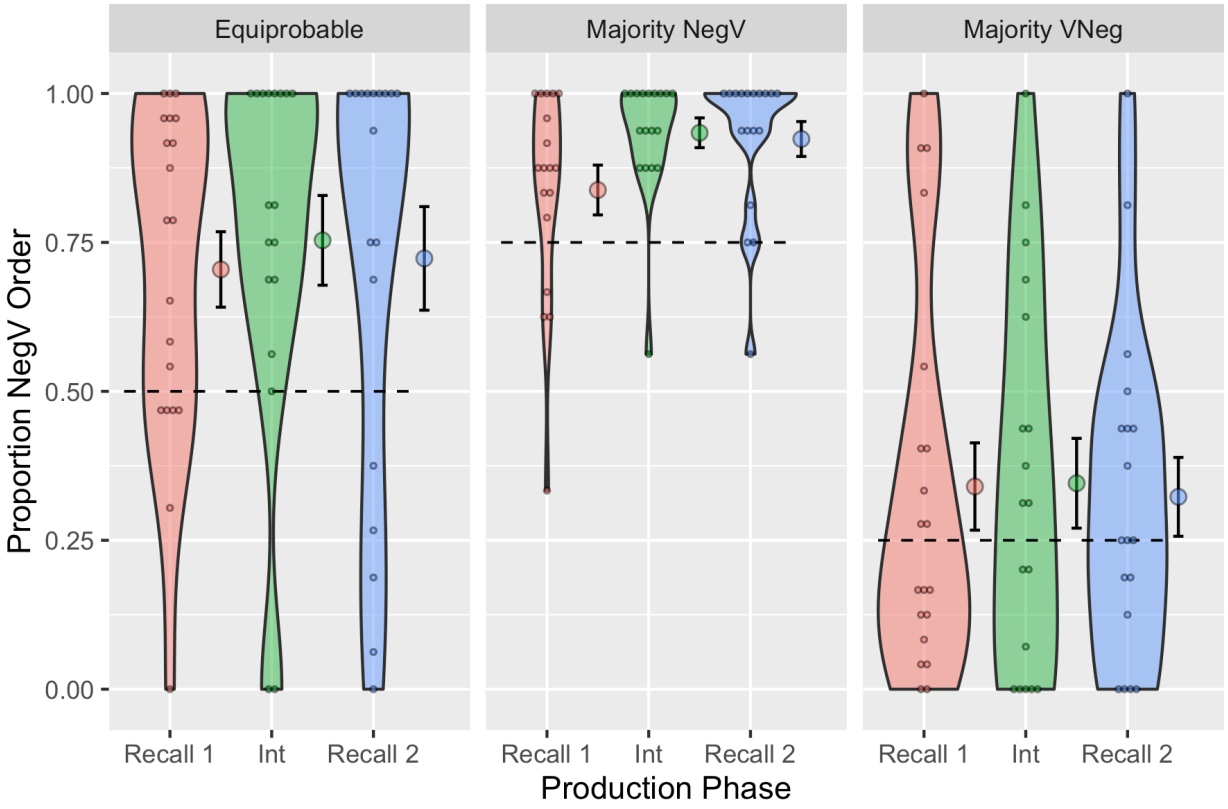
[10] A more strict measure of accuracy which required exact matches was used in the analysis of this experiment in Burgess (2022); the overall pattern of results does not differ by using this more lenient accuracy measurement.

accuracy and production accuracy models included Phase, Condition, and their interaction as fixed effects. For each, backwards stepwise elimination identified the model with only a fixed effect for Phase as the best fit. Notably, including Condition as a fixed effect did not result in a better fit, indicating that the input condition was not an important factor in explaining accuracy in either task.

Neg-First Bias. Recall that a Neg-First bias predicts that the strength of regularization would vary between the Majority NegV and Majority VNeg Conditions, with participants in the Majority NegV Condition regularizing the majority word order more. We also might expect that if there is a strong bias towards preverbal negation, regularization in the Equiprobable Condition would be in the direction of preverbal negation. Figure 6 shows the proportion of NegV order participants produced in each condition and critical production stage. Visual inspection of the data shows that across conditions, participants produced more preverbal negation on average compared to the proportion of preverbal negation in their training input; articulated in terms of use of the *majority order*, participants in the Majority NegV Condition appear to use the majority order more than those in the Majority VNeg Condition.

These effects were analyzed using mixed effects logistic regression. First, to test whether participants in the Equiprobable condition produced NegV order significantly greater than chance, I ran a model including NegV Order as the binary dependent variable, fixed effects of Condition (ref = Equiprobable), Phase (ref = Recall 1), and their interaction, and random effects of participant nested in group. This model showed a significant positive intercept ($\beta = 1.53$, SE = 0.50, $p = 0.002$), indicating that the choice of NegV order in the Equiprobable Condition was significantly greater than chance in the Recall 1 phase of the Equiprobable Condition.

Figure 6. Proportion of NegV order production in Experiment 2 for all conditions and critical production stages. Dashed lines indicate proportion of preverbal negation order in the input; small points show subject means; the large points offset to the right of each violin plot indicate the overall mean for each condition; error bars indicate standard error of participant means.



To test whether participants in the Majority NegV Condition were more likely to boost the majority order than participants in the Majority VNeg Condition, I ran a model including Majority Order as the binary dependent variable, fixed effects of Condition (Helmert coded to test Majority VNeg vs. Majority NegV, and Equiprobable vs. combined mean of other two conditions), Phase (Helmert coded to test Interaction vs. Recall 1, and Recall 2 vs. combined mean of the other two conditions), and their interaction, and random effects of participant nested in group. Note that for the Equiprobable Condition, NegV was arbitrarily chosen to represent the majority order. This model revealed a significant main effect indicating that the use of the majority order was significantly lower in the Majority VNeg Condition compared to the Majority

NegV Condition ($\beta = -1.03$, SE = 0.35, $p = 0.003$). There was also a main effect of phase indicating that the majority word order was boosted more in the Interaction Phase than the Recall 1 Phase ($\beta = 0.24$, SE = 0.07, $p < 0.001$), though note that choice to include NegV as the majority order for the Equiprobable Condition contributes to this result. The model also revealed a significant interaction term, indicating that boosting of the majority order in the Interaction Stage compared to Recall 1 was smaller in the Majority VNeg Condition compared to the Majority NegV Condition ($\beta = -0.34$, SE = 0.09, $p < 0.001$).

Effect of Interactive Task. To test whether participants produced a greater proportion of NegV word order in Interaction compared to Recall 1, pairwise comparisons were conducted on the previously described model with Tukey method p-value adjustment using the *emmeans* package for R (Lenth, 2020). These comparisons revealed that within each condition, the only significant differences in use of the NegV order between Production Phases were within the Majority NegV Condition; within this condition, there was significantly more NegV use in the Interaction Stage than the Recall 1 stage ($\beta = -1.20$, SE = 0.28, $p < 0.001$), and NegV use was also significantly greater in Recall 2 compared to Recall 1 ($\beta = -1.02$, SE = 0.27, $p < 0.001$) indicating a lasting effect of interaction.

4.3.3 Discussion

Experiment 2 demonstrates the existence of a behavioral bias consistent with Neg-First preferences in adult English speakers learning and producing a language which contained probabilistic variation of both preverbal and postverbal negation. Participants in the Equiprobable and Majority NegV produced significantly more NegV word order in their productions than they were exposed to in training, and those in the Majority VNeg condition did not regularize towards using more VNeg order, but instead produced numerically greater NegV

order. This finding is consistent with observations of Neg-First preferences in naturalistic language learning settings, such as the tendency for beginning and intermediate L1 English learners of Korean to produce the Korean preverbal negation construction in greater proportions than L1 Korean-speaking adults, compared to the roughly synonymous postverbal negation construction (Kim & Yun 2013). Experiment 2 shows that this Neg-First preference is still present even when variant frequency and grammatical complexity are controlled for. Whether this preference depends on prior language experience is not something addressed in the present study. However, the exact predictions of English-language transfer are unclear. Because English negation always follows a finite auxiliary verb, but precedes a lexical verb, English speakers may be biased to mark negation prior to the lexical verb simply on the basis of their language experience. To address the question of whether this preference is dependent on language experience, Experiment 3 expands this research to speakers of Japanese, an SOV language in which negative marking follows the lexical verb but precedes tense marking.

The study also investigated whether communication played a role in encouraging a preference for earlier negation marking, finding that *only* participants within the Majority NegV Condition used significantly more preverbal negation order during (and after) interaction compared to production in isolation prior to interaction. This result is not consistent with the idea that a communicative context induces a greater preference for earlier negation in general. However, the presence of an effect in the NegV Condition, in which the majority order and Neg-First bias were aligned, may be consistent with previous research indicating that communicative interaction gives rise to increased regularization via mechanisms of reciprocal priming and/or asymmetric accommodation (Fehér, Wonnacott, & Smith, 2016; Fehér, Ritt, & Smith, 2019).

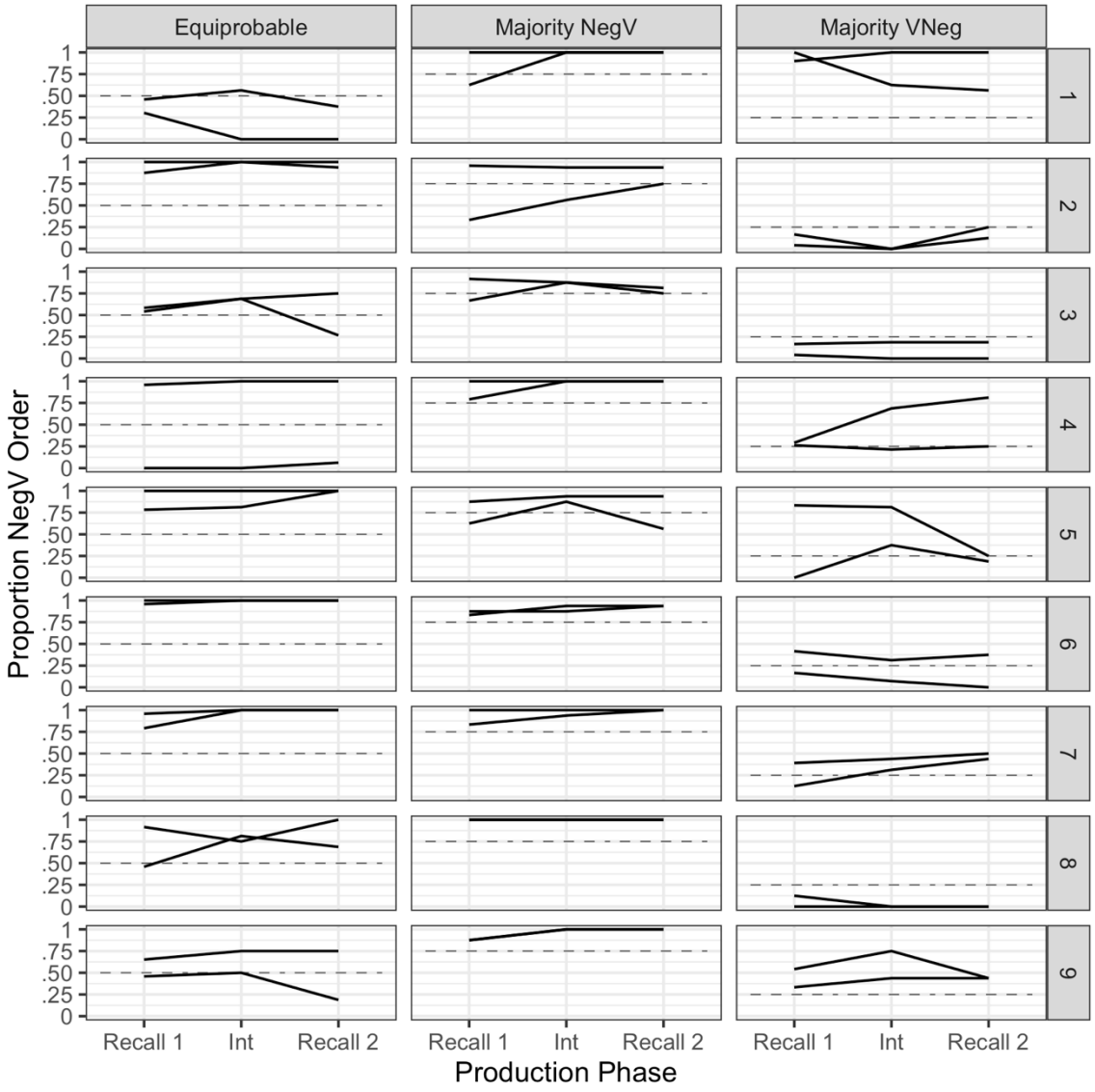
Patterns of priming, alignment, and regularization within dyads are explored further in the following section.

4.3.4 Exploratory analysis: behavior within dyads

Consider that one possible explanation for a lack of increased regularization during interaction is that, if the baseline regularization by participants producing the language during isolation is already high, it may not increase further when communication is added (c.f. Saldana et al. 2021). However, this is not the case here, since the only significant increase in regularization occurred in the Majority NegV Condition, which was the condition which showed the greatest amount of regularization in the Recall 1 phase (in the direction of more consistent NegV order). In this case, high (but not ceiling-level) regularization at the individual level in the Recall 1 phase seemed to encourage more regularization during interaction. Structural priming, in which interlocutors match the syntactic structure of their partners' utterances, could be one mechanism which influences participants to converge during interaction (Pickering & Garrod 2004). In the Majority NegV Condition, reciprocal structural priming would lead to more regularization since participants' biases towards NegV order are reinforced during communication. This is consistent with previous research suggesting that reciprocal priming encourages interacting pairs to converge on more regular systems (Fehér et al. 2016). Another mechanism which might be at play here is asymmetric accommodation: variable users are more likely to accommodate categorical partners in interaction than vice versa (Fehér et al. 2019). To determine whether these mechanisms could plausibly play a role in determining the pattern of results in which a reliable change in the amount of preverbal negation order from Recall 1 to Interaction only occurred within the Majority NegV Condition, I take a closer look at how pairs behaved in terms of priming, alignment, and regularization across conditions and phases.

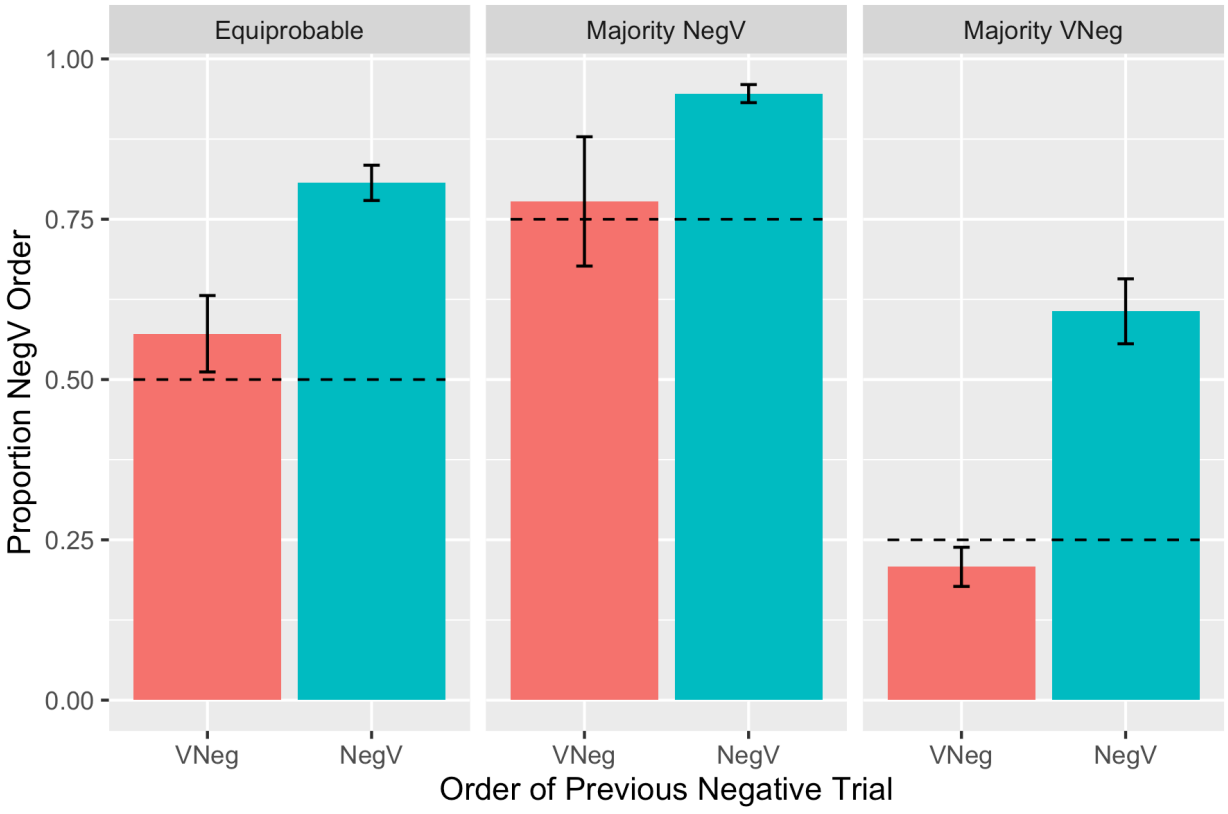
Figure 6 shows participants' proportion of NegV order use across conditions and phases in separate by-pair plots. In this figure, lines which overlap indicate that the players were aligned (i.e., used similar proportions of preverbal negation order). Impressionistically, we see that many pairs of participants in the Majority NegV Condition and Equiprobable Condition converged towards highly regular or even categorical NegV order use, while convergence towards highly regular VNeg use was less common, even in the Majority VNeg Condition.

Figure 6. Proportion of NegV order production in Experiment 2 with separate plots for each interaction pair. Dashed lines indicate proportion of preverbal negation order in the input.



Priming of word order use during interaction. To investigate whether structural priming (i.e., the tendency to match the order of the previous utterance) played a role in interaction, we can test whether word order choice was influenced by the word order of the previous negative trial. Figure 7 shows, for each condition, the proportion of trials that participants produced NegV order, based on the word order of the preceding negative trial (note that one third of trials during interaction were randomly interspersed affirmative sentence fillers, so the preceding negative trial was not always the *immediately* preceding trial). The data indicate priming in all three conditions: participants were more likely to use NegV order if the preceding negative trial produced a sentence in NegV order.

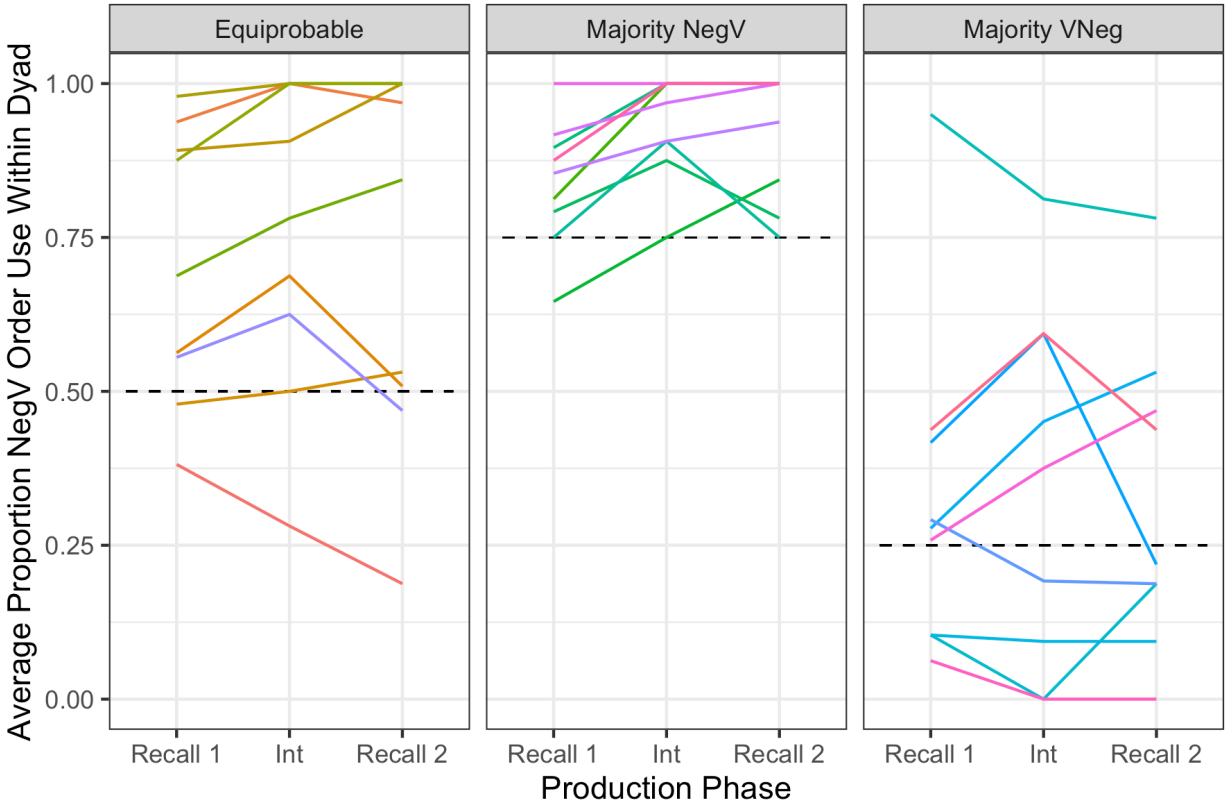
Figure 7. Proportion of sentences produced in NegV order during interaction based on the word order of the preceding negative trial. Dashed lines indicate proportion of NegV order in the training input; error bars indicate standard error.



This impression is confirmed by logistic mixed-effects regression model with NegV order use as the dependent variable, Condition and Previous Negative Trial Order as fixed effects, and participant nested in group as a random intercept. Both fixed effects were Helmert coded. This model revealed that participants were significantly more likely to use NegV order when NegV order was used on the previous trial ($\beta = 0.75$, SE = 0.33, $p = 0.02$). Additionally, the model revealed the expected findings that NegV order was used significantly more than chance overall, based on a significant positive intercept ($\beta = 1.63$, SE = 0.51, $p = 0.001$), and that NegV order was used significantly less in the VNeg Condition than the NegV Condition ($\beta = -2.66$, SE = 0.67, $p < 0.001$). The model revealed no significant interactions between Condition and Previous Trial Order.

Combined system regularity. Although there is evidence of structural priming in all three input conditions, note that reciprocal priming may not induce a strong regularization effect if the combined system of the two interlocutors is highly variable. For example, if participants within a dyad regularize the input in different directions, then the combination of the two systems will still result in high variability. In this case, a bias to match the partner (i.e., a priming bias) and a bias to maximize the most frequently used word order (i.e., a regularization bias) may compete against each other. Moreover, for participants in the Majority VNeg Condition, a bias to maximize the most frequently used word order will compete with a Neg-First bias even prior to interaction, leading to more variable combined systems. Figure 8, which shows the average proportion of NegV word order use for each dyad, shows a stark contrast in the spread of variability of the word order in the combined systems of dyads before, during, and after interaction among those in the Majority NegV Condition compared to the other two training conditions.

Figure 8. Average proportion of NegV order use by dyad across conditions and phases. Each colored line represents a dyad. Dashed lines indicate proportion of NegV order in the training input.



Another mechanism which has previously been shown to operate during interaction, and which may play a role in more dyads converging on regular systems in the Majority NegV Condition compared to the other two conditions, is asymmetric accommodation: variable partners are more likely to adjust their production to align with categorical partners than vice versa (Fehér et al. 2019). This is of relevance in the present study because among the 11 participants that used an entirely categorical system in the pre-interaction Recall 1 phase, more favored NegV order than VNeg order. 8 participants used NegV order categorically (3 in the Equiprobable Condition, 5 in the Majority NegV Condition, 1 in the Majority VNeg Condition) whereas only 3 participants used VNeg order categorically (1 in the Equiprobable Condition, 2 in the Majority VNeg Condition). Variable users may be expected to accommodate to categorical

users due to asymmetric accommodation, and categorical users were more likely to favor NegV order, leading to more dyads converging on NegV order during interaction.

Due to mechanisms of priming and asymmetric accommodation, we should expect variable users who interact with a categorical partner to shift to a more regular system, in comparison to variable participants who interact with another variable partner. Regularity of a system can be measured using Shannon entropy given by the formula $-\sum p(x) \log p(x)$ where the sum is over the two possible word orders (NegV or VNeg) and *p(x)* is the frequency of word order *x* in a participant's productions. With two possible word orders, entropy can range from 0 to 1: an entropy of 0 corresponds to a participant who consistently uses a single word order, and an entropy of 1 corresponds to a participant who uses both word orders in equal proportion.

To test the hypothesis that variable users who interact with a categorical partner are more likely to show a decrease in entropy than those who interact with another variable partner, each participant was coded for two binary factors: whether they were variable or not in the Recall 1 phase, and whether their *partner* was variable or not in the Interaction phase. These factors were then used to generate a linear mixed-effects model predicting entropy of users who were variable during Recall 1. The model included fixed effects of Condition (ref = Majority NegV), Phase (ref = Recall 1), Categorical Partner (ref = non-categorical partner), and interaction of Phase and Categorical Partner, and a random intercept for Group. P-values for this model were calculated using the lmerTest package using Satterthwaite approximation (Kuznetsova et al. 2017). The model revealed significant interactions that confirm that presence of a categorical partner at interaction led to a decrease in entropy over the course of the experiment: when variable users encountered a categorical partner at interaction (as opposed to a variable partner), the entropy of

their productions decreased more from Recall 1 to Interaction ($\beta = -0.30$, SE = 0.12, $p = 0.02$) as well as from Recall 1 to Recall 2 ($\beta = -0.26$, SE = 0.12, $p = 0.04$).

Response Times. Given the assumption that postverbal negation incurs processing difficulty, one may reasonably hypothesize that, on average, participants in the matcher role would take longer to respond to postverbal negation than preverbal negation, and/or that participants in the director role would take longer to produce postverbal negation than preverbal negation. The response times for matchers and directors in the interaction phase are shown in Table 6. While it is difficult to draw strong conclusions from this data because the task was not designed with explicit time pressure, a linear mixed-effects regression model with the (matcher's) selection response time as the dependent variable, Condition, NegV Order, and their interaction as fixed effects, and participant nested in group as random effects did not show any significant effects. Another model with the same fixed and random effects, but with the (director's) production response time as the dependent variable, showed a marginal main effect of NegV Order, such that directors, on average produced sentences with NegV order slightly faster than those with VNeg order ($\beta = -1.43$, SE = 0.77, $p = 0.06$). These results point towards producer biases more so than comprehension biases favoring NegV order.

Table 6. Average response time (in seconds) by condition and negation order for director and matcher trials (standard error in parentheses).

Director

	Equiprobable	**Majority NegV**	**Majority VNeg**
NegV	12.3 (0.4)	14.3 (0.4)	15.3 (0.9)
VNeg	16.8 (1.1)	16.2 (1.9)	13.6 (0.6)

Matcher

NegV	8.8 (0.6)	10.1 (0.3)	11.1 (0.6)
VNeg	9.0 (0.3)	9.7 (0.9)	10.1 (0.5)

4.3.5 Participant Language Background

Information about participants' language experience was collected through a post-experimental questionnaire hosted on Qualtrics. The questionnaire asked participants to name up to five significant languages (including English) that they used or studied currently or in the past, and presented questions about age of acquisition, fluency in using and understanding the language, and frequency and types of exposure to the language (see Appendix B for the full list of questionnaire questions). Besides English, the languages that participants most frequently identified having used or studied were Spanish (n=33), French (n=17), German (n=10), and (Mandarin) Chinese (n=8). Though it would be interesting to examine whether participants with greater experience or fluency with postverbal negation languages showed different behavior than those without such experience, this study was not designed to investigate this question adequately. Few participants had extensive experience (> 5 years total study/use) with languages which use postverbal negation (German (n=5), Malayalam (n=1), Korean (n=3). The diversity of these languages and the variety of negation constructions within them also make it difficult to predict how language transfer would operate, especially since both German and Korean also use preverbal negation constructions. Therefore, I leave more rigorous investigation of the effect of language transfer on biases to use preverbal negation for future work. This question is also taken up further in Experiment 3 in Chapter 5 of this dissertation.

4.3.6 Comparison to Experiment 1

Why did the participants in Experiment 2 show a significant Neg-First bias, even prior to interaction, whereas the participants in Experiment 1 did not? In addition to the presence of an interactive director-matcher task, the design of Experiment 2 differed from Experiment 1 in several ways. Whereas Experiment 1 included a word bank during the production phases to help

participants remember lexical items, this word bank was not present in Experiment 2. Since participants in Experiment 1 did not show a strong tendency to regularize overall, this change was made to increase the difficulty of lexical retrieval, as increased demands on lexical retrieval have been shown to increase regularization of unconditioned variation in adult learners (Hudson Kam & Chang 2009). Thus, the removal of the word bank may have contributed to the increased regularization seen in Experiment 2 as intended.

Another difference between Experiment 1 and Experiment 2 was the addition of a negative sentence comprehension task that required the participant to apply truth-functional negation to choose the correct image (e.g., when given a sentence corresponding to "not punch boxer artist", the participant should *avoid* choosing the image in which the boxer is punching the artist). The possible impact of this change is far more speculative. If increased awareness or activation of the semantic contribution of negation is significant to generating a Neg-First bias, either because it increases implicit or explicit connections by the learner to the patterns of negation they have encountered in their previous language experience, or because the function of negation is particularly important or salient leading to earlier production, then the reinforcement of the function of the negative particle in this task may have encouraged a larger Neg-First bias in Experiment 2.

Finally, Experiment 2 used VSO word order instead of SVO word order. Therefore, in Experiment 2, the preverbal negation order also allowed negation to be sentence initial, while this was not permitted in the language for Experiment 1. It may be the case that NegV order was more greatly preferred in Experiment 2 because this led to not only preverbal negation, but also sentence-initial negation. Another piece of evidence that makes this hypothesis worth considering is that children acquiring a language with clause-internal negation often go through

an early stage of using clause-peripheral negation (see §2.1.3). A preference for clause-initial negation could trade-off with a preference for surface scope transparency; in sentential negation, the subject is not usually considered to be within the scope of negation. This is illustrated by the famous sentence *The king of France is not bald*, which is usually taken to convey the understanding (the presupposition) that the king of France exists. Thus, many formal theories of negation incorporate this understanding by arguing or assuming that negation is internal and does not have scope over the subject (e.g., Frege 1892, Strawson 1950). Hence, sentence-initial negation (like NegVSO) allows negation to be as-early-as-possible in the sentence, but subsequent material (such as the subject) is not all captured in the scope of negation, leading to reduced scope-transparency. Future artificial language learning research can examine to what degree and in what circumstances clause-initial or clause-peripheral negation is preferred over (merely) preverbal negation by manipulating word order possibilities and negative scope interpretations of the artificial language.

4.4 Conclusion

Though the cross-linguistic tendency for the standard negator to precede the verb has long been noted by typologists, little is known about the source of this tendency. Experiment 2 provided a first controlled investigation to the hypothesis that audience design plays a role in motivating a Neg-First bias in communication by comparing Neg-First biases in isolated production and in an interaction task.

The interaction task was found to increase NegV use significantly above that in the prior Recall 1 task *only* in the Majority NegV Condition. This result appears to be consistent with prior research indicating that mechanisms of structural priming and asymmetric accommodation lead to dyads which converge on regular systems, particularly in the presence of a strong bias

towards a specific pattern (Fehér, Ritt, & Smith, 2019; Fehér, Wonnacott, & Smith, 2016; see also Saldana, Smith, Kirby, & Culbertson, 2021 for n.s. effect of interaction). However, it is not consistent with the idea that a communicative context induces a greater preference for earlier negation in general. Therefore, it is worth considering other potential motivations for a Neg-First tendency in learning and typology, such as those which center production difficulty à la the Production- Distribution-Comprehension approach (MacDonald, 2013). For example, the possible conceptual salience of negation could provide a general explanation for negation being accessed and produced earlier. Language transfer and the preference to produce word orders that are congruent in the L1 and L2 could provide a language-experience-dependent, producer-centered explanation for Neg-First biases when the L1 has preverbal negation (as in English).

The present study did find that adult English speakers showed a Neg-First bias in production even prior to interaction. This is consistent with previous studies finding regularization in the absence of interaction when certain variants are preferred (Culbertson et al. 2012, Hudson Kam & Newport 2009). However, it may be the case that this is driven by L1 transfer effects, since in English, the standard negator *not* always precedes the main verb. The role of prior language experience is tested in the next chapter by expanding this research to speakers of Japanese, in which the standard negator follows the main verb.

Chapter 5 What Role Does Previous Language Experience Play in Neg-First Biases?

This chapter considers the extent to which previous language experience explains tendencies for second language learners to produce preverbal negation in second language acquisition. I conduct a close inspection of previous studies which have examined the influence of first language (L1) on the acquisition of negation in second language (L2) acquisition in both artificial and naturalistic language learning settings. I then introduce Experiment 3, an artificial language learning study investigating whether speakers of a postverbal negation language (Japanese) show a preference for preverbal negation when learning a miniature artificial language which contains both preverbal and postverbal negation.

5.1 Background and Motivation

Neg-First tendencies observed in natural language acquisition and typology have sometimes been claimed to result from universal cognitive mechanisms at play in language acquisition and use (e.g., Plag 2008, Horn 1989). Though Experiment 2, reported in Chapter 4, confirmed that English speakers learning a miniature artificial language which contained both preverbal and postverbal negation showed a bias to overproduce the preverbal negation construction, these results do not verify the existence of a non-language specific bias, since English is a language in which the standard negator *not* always appears before the main verb (e.g., 'I do *not like* tomatoes'). The bias for participants to overproduce preverbal negation in Experiment 2 may merely reflect prior experience with a preverbal negation language; increased use of syntactic structures that are shared between languages is a well-known phenomenon in acquisition and contact-induced language change (e.g., Silva-Corvalán 2008, Toribio 2004). It is

important to further test the existence of a preverbal negation preference with speakers of a postverbal negation language, because we should be cautious about overgeneralizing language behavior and biases based on evidence drawn only from narrow populations in terms of language background. For example, Martin & Culbertson (2020) found evidence challenging a previously proposed universal perceptual preference for suffixing, which had been primarily based on experiments with English speakers, finding that speakers of Kîîtharaka, a heavily prefixing Bantu language spoken in rural Kenya, had opposite judgements to English speakers.

If a bias in learning is shown to occur independently of language experience, such that speakers of languages which violate a given pattern are still shown to favor that pattern in learning, this provides stronger evidence that the bias is a plausible candidate for shaping linguistic typology. Several experimental studies using artificial language learning and silent gesture paradigms have identified preferences in learning or production which can be linked to typological tendencies of word and morphology order and are found among participants of different language background groups (Motamedi et al. 2022, Saldana, Oseki, & Culbertson 2021, Culbertson et al. 2020). In a series of artificial language learning studies testing the effects of previous language experience on biases related to word order during learning, Culbertson et al. (2020) tested French- and Hebrew-speaking children and adults' word order preferences for adjective and numeral nominal modifiers. Though French and Hebrew place adjectives and numerals on opposite sides of the noun (N-Adj, Num-N), Culbertson et al. (2020) found a preference for participants to produce adjectives and numerals on the same side of the noun, which parallels tendencies of nominal phrase order in typology. However, even when overall preferences are observed among learners with different language backgrounds, differences between groups reveal a complex interplay between L1, L2, and general cognitive biases which

needs to be considered. Culbertson et al. (2020) found that French- and Hebrew-speaking adult learners who were L2-speakers of English showed a preference for English-like pre-nominal harmony over the other patterns, and there was also some evidence that English learners who had L2-experience with languages that used (N-Adj, Num-N) nominal ordering used that order more often than a comparable population of English monolinguals. Culbertson et al. (2020) conclude that this is consistent with hypotheses that the L2 can take on a stronger role than the L1 as a source of language transfer while learning subsequent languages (Bardel & Falk 2012), while suggesting that the degree of L2 influence may be modulated by general learnability, with more easily learnable patterns more likely to influence the bilingual speaker.

Arguments for an experience-independent Neg-First bias in acquisition are largely based in previous research of L2 acquisition of negation, which has shown that even when both the L1 and the L2 contain postverbal negation, learners often go through stages in which they overproduce preverbal negation or systematically produce sentences such that the negator appears earlier in the sentence than it would in the L2, as described in §2.1.3. These tendencies occur particularly when the target language has variable placement of negation (e.g., postverbal placement in main clauses, but preverbal placement in subordinate clauses). However, it is not clear whether these patterns reflect a preference for preverbal negation *per se* instead of a preference for other features which factor into the choice to use preverbal negation. To clarify such ambiguity, I will review two studies, Hyltenstam (1977) and Kim & Yun (2013), which examined the influence of L1 on learning and developmental sequences of negation, particularly in cases where postverbal negation is present in both L1 and the L2.

Hyltenstam (1977) reports developmental sequences in the acquisition of Swedish negation. In Swedish, the negator *inte* is placed immediately after the finite verb in main clauses

(e.g., Kalle *kommer inte* idag 'Kalle comes not today') and immediately before the finite verb in subordinate clauses. (e.g., Det är skönt att Kalle *inte kommer* idag 'It's fine that Kalle not comes today'). Based on a written production test taken by 160 students studying Swedish as a foreign language at a Swedish University, with 35 different L1 backgrounds represented, Hyltenstsam (1977) found a similar route of acquisition for all learners regardless of language background. Specifically, in main clauses, participants were more likely to follow Swedish norms (i.e. placement after the finite verb) with finite auxiliary verbs (e.g., *kan* 'can') than with finite main verbs (*kommer* 'come'). In subordinate clauses, however, among participants who differentiated clause types at all, the reverse was true; participants were more likely to follow Swedish norms (i.e., placement before the finite verb) with finite main verbs than with auxiliary main verbs. Therefore, preverbal negation was more often favored with main verbs than with auxiliary verbs, regardless of participants L1 language background. Notably, this same pattern was observed among participants who spoke Turkish as an L1 even though Turkish expresses negation as a suffix on the main verb, consistent with a Neg-First bias in learning independent of language background.

However, whether this learning pattern truly reflects a preference for preverbal over postverbal negation is contestable. That learners show different outcomes with auxiliary verbs and main verbs is also well-attested in language acquisition of negation in numerous other target language where negation follows the finite verb (e.g., English: Giuliano 2003, Wode 1981; French: Giuliano 2003, Meisel 1997; German: Dimroth 2008, Meisel 1997; see Table 3 in §2.1.3). In these languages, the acquisition of negation is tied up in finite verbal inflection. Talking about child language acquisition, but applicable also to adult L2 acquisition, Dimroth (2010: 58) argues that "this variation [in negator placement with different verb types] is not due

to problems with the placement of negation but rather reflects different stages in children's acquisition of finiteness." In this line of theorization, correct placement of negators depends on the acquisition of finiteness, and the category of finiteness is easier to grasp for L2 learners when carried by non-thematic verbs (Dimroth 2010). Another reasonable possibility for the differences in negator placement with different verb types, raised by Hyltenstam (1977), is that the correct (postverbal) placement of negation relative to auxiliary verbs for Swedish is learned more quickly because auxiliary verbs are more frequently encountered, according to Swedish corpora.

Next, we will turn to a study that tested the influence of L1 on the learning of negation in Korean. Whereas Swedish contains both preverbal and postverbal expression of standard negation but each order is relatively constant within in a given clause type, Korean uses both a postverbal negation construction (example 30a) and preverbal negation construction (example 30b) used in main clauses.

(30) Korean (Chang 1996: 101, 103)

a. *Yong-i TV-lul po-ci* ***an****-ha-yo.*
Yong-SBJ TV-OBJ see-SUSP NEG-AUX-POL
'Yong doesn't watch TV'

b. *Yong-un TV-lul* ***an*** *pwa-yo.*
Yong-TOP TV-OBJ NEG see-POL
'Yong doesn't watch TV.'

Kim & Yun (2013) investigate the learning and developmental sequence of Korean negation forms among English and Japanese learners of Korean with a written production test. Kim & Yun (2013) found that both L1 English and L1 Japanese learners of Korean produced more of the preverbal negation form than the postverbal negation form, and both produced the preverbal negation form significantly more than a control group of Korean-speaking adults.[11] However,

[11] Kim & Yun (2013) interpret their statistical model as showing that the Japanese L1 group produced more long-form (postverbal) negation than the English L1 group did, but the raw numbers provided in the article suggest the opposite (perhaps unexpectedly given that Japanese uses postverbal negation). The raw data reported show that Korean adults produced postverbal negation in 75-84% of trials among the different sentence type conditions

this preference for the preverbal negation construction among learners of Korean, even among speakers of a postverbal negation language like Japanese, does not necessarily indicate that learners are showing a preference for preverbal negation *per se.* Firstly, we have little information about the nature of the language input that learners in this study received. If the proportions of preverbal and postverbal negation constructions that are presented in instructional materials and methods differ greatly from those found in Korean outside of an instructional context, it could be the case that learners are mimicking the input they have been exposed to through instruction, even though they are diverging from the Korean adult norm. Secondly, the Korean preverbal and postverbal negation constructions also differ in that the preverbal negation construction uses a particle *an* whereas the postverbal negation constructions involves a suffix -*ci* followed by a negative verb *anhta* which needs to be conjugated. Learners may find the particle negation construction relatively simpler or easier, for reasons unrelated to the fact that it occurs before the verb, because it only requires adding a particle to the left of the verb found in a corresponding affirmative sentence. On the other hand, as Kim & Yun (2013) note, it is not necessarily obvious that this should be the simpler option, especially for certain types of verbs. For example, in light verb constructions which combine an object and light verb *hada* 'to do' such as *kwongbu-hada* 'to study', the negative particle intervenes between the noun and light verb *kwogpu an hada* 'to not study', and learners must know to separate the noun and light verb in this case.

In Experiment 3, I control the potentially confounding variables of morphosyntactic complexity and learner input frequencies (both the frequency of each negation construction that

reported in the article, English L1 learners produced postverbal negation in 45-56% of trials, and Japanese L1 learners produced postverbal negation in 17-38% of trials. I am basing my discussion off of the raw numbers and the overall discussion and conclusions of the authors.

the learner sees and the collocational frequencies of each construction) by using an artificial language learning paradigm.

5.2 Experiment 3: Investigating Neg-First biases among Japanese speakers

To test whether a bias to produce preverbal negation exists among people learning a language that freely allows both preverbal and postverbal negation, even when learners have extensive language experience with a postverbal negation language, Experiment 3 is a partial replication of Experiment 2 with Japanese-speaking participants. In Japanese, standard negation is expressed as a suffix on the verb, appearing after the main verb and before affixes expressing tense, as shown in example (31).

(31) Japanese
Taroo-wa asagohan-o tabe-na-katta.
Taroo-TOP breakfast-OBJ eat-NEG-PST
'Taroo didn't eat breakfast.'

The design and procedure for Experiment 3, described fully in the following section, is identical to Experiment 2 with two major differences. First, the lexicon for Experiment 3 was adapted by a Japanese speaker to conform to Japanese phonotactics and avoid Japanese homophones. Second, Experiment 3 does not include a dyadic interaction component following the training and final solo production task.

5.2.1 Methods

Participants. 67 Japanese-speakers were recruited in Japan through word-of-mouth and classroom visits at the University of Tokyo. All participants grew up speaking Japanese at home and/or in school, lived in Japan at the time of the study were recruited, and completed the study remotely on a computer. Data from participants that failed to meet an accuracy threshold of 75% in the final production phase were excluded from analysis. Based on this criterion, data from 7

participants were excluded from analysis, such that 60 participants were included in analysis with 20 participants per condition. Participants were compensated with a ¥2000 gift card to Amazon.jp upon completion of the study.

Materials. Lexical items in the language were displayed in katakana script; this script is used for the transcription of foreign words into Japanese. As in Experiment 2, the artificial language included six nouns, four verbs, and a negation particle (see Table 7) and used VSO word order with the negative particle able to appear either immediately before or after the verb. Whether a sentence appeared with pre- or postverbal negation ordering was not conditioned on any aspect of the sentence itself, but the proportion of each negation order in the input language varied depending on the input condition the participant was assigned to. Audio of these sentences were synthetically generated using Amazon Polly's 'Mizuki' voice[12] and each sentence was paired with the same cartoon images as those used in Experiments 1 and 2.

Table 7. Artificial language lexicon for Experiment 3 with broad IPA transcriptions

Nouns	**Exp 2**	**Exp 3**	**Verbs**	**Exp 2**	**Exp 3**
boxer	daki	ドキ /doki/	shoot	patu	パツ /patsɯ/
chef	falit	ホレト /hoɾeto/	kick	sal	セル /seɾɯ/
burglar	ludin	ルピン /ɾɯpin/	point	toma	トマ /toma/
cowboy	mook	ムク /mɯkɯ/	punch	umi	プミ /pɯmi/
doctor	pepo	ピポ /pipo/	**Negation Particle**		
artist	sido	シド /ɕido/	'not'	pik	ピク /pikɯ/

Design & Predictions. The experimental design is the same as Experiments 1 & 2. Participants were randomly assigned to one of three language input conditions in which the proportion of preverbal and postverbal negation was manipulated: participants in the Majority

[12] https://aws.amazon.com/polly/

NegV Condition were exposed to majority (75%) preverbal negation ordering, participants in the Majority VNeg Condition were exposed to majority (75%) postverbal negation ordering, and participants in the Equiprobable Condition were exposed to 50% of each ordering. A bias in favor of preverbal negation predicts that participants in the Majority NegV Condition will regularize towards NegV word order to a greater extent than participants in the Majority VNeg Condition will regularize towards VNeg word order. We may also expect regularization in the Equiprobable Condition would be in the direction of overproducing preverbal negation, if regularization occurs at all.

Procedure. Each participant was trained and tested on the artificial language in a single session no more than 60 minutes long. The experiment was created using PsychoPy (Pierce et al. 2019) and hosted on the Pavlovia platform.[13] Participants accessed the online study using their own computers; participants were emailed a link to an informed consent form which automatically redirected them to the experiment on Pavlovia when they consented to participating in the study and submitted the form. Upon completion of the experiment, they were automatically directed to take a questionnaire hosted on Qualtrics which included questions about language background (see Appendix B). Participants worked through series of three training stages in which they were exposed to the artificial language, which were identical to those in Experiment 2, outside of the lexicon and instructions being presented in Japanese.

Stage 1, Noun Training. Participants were introduced to the six nouns of the artificial language over 18 exposure trials. Then, participants were tested on their comprehension and production of these nouns with 12 trials of a two-alternative forced choice task (2AFC) followed by 12 production trials.

[13] https://pavlovia.org/

During this phase, and all training blocks to follow, the trials were presented as follows. On each exposure trial, participants were presented with an image and a corresponding description in the language, visually printed beneath the image and aurally presented. On each 2AFC trial, two images were presented on the screen above a description, and participants were directed to use the left and right arrow keys to choose the image that corresponded to the description. The correct answer was on the left side half of the time. Participants received feedback on their responses. For production trials, an image was presented in the center of the screen and participants were asked to type the text that described the image into a text box. After participants submitted their answer, the correct answer appeared on the screen for 3 seconds regardless of participant accuracy.

Stage 2, Sentence Training. Participants were introduced to the language's four verbs in the context of complete transitive affirmative sentence over 48 exposure trials. Sentences were balanced so that each noun was presented equally often in both subject and object position with each verb. This was followed by 12 2AFC trials and 12 production trials not yet encountered by participants during the exposure stage. The distractor images in the 2AFC trials differed from the correct image by a single element (agent, action, or patient).

Stage 3, Negative Sentence Training. Next, participants were exposed to 48 negative sentences in the language. The images used in this phase resembles those images used for affirmative sentences, with a slightly transparent red circle with a diagonal line superimposed over the center of the image. As mentioned above, the proportion of preverbal and postverbal negation varied between experimental conditions.

Participants completed 12 2AFC to test comprehension of negative sentences. For these trials, one of the images depicted the negated event, making this the *incorrect choice*, and the

other image differed in the depiction of the patient. Participants received feedback on this task as in previous tasks. If they chose the incorrect image, they received a reminder that "negative sentences tell you what did not happen". By having participants choose between 'affirmative' event drawings to test negative sentence comprehension, we can check that participants understood the function of the negation marker without relying on the 'negation symbol' cue in the drawing.

The production phase following negative sentence exposure and comprehension trials comprised 36 totals trials using images withheld from previous stages of the experiment (24 negative sentences and 12 affirmative sentences). Feedback was not provided to participants during this stage.

5.2.2 Results

All statistical analyses were conducted in R, and mixed effects logistic regression models were fit using the lme4 package (R Core Team 2021, Bates et al. 2015). The data files and analysis code are available at https://github.com/DaniLBurgess/BurgessDissCh4Ch5.

Prior to production response analysis, typos were automatically corrected by splitting each response typed by a participant into a series of words by splitting the string at the spaces and identifying the closest match for each among the artificial language lexicon. Levenshtein string edit distance was used to determine the closest match. Levenshtein distance is the minimal number of insertions, deletions, and symbol substitutions required to transform a string to another. If a participant's word was one Levenshtein edit distance away from only one legal word, then the participant's response was corrected to that legal word. Then, the word order for negative sentence trials (NegV vs. VNeg) was automatically labeled based on the closest match when comparing the whole corrected response to each of the correct target sentences. Full

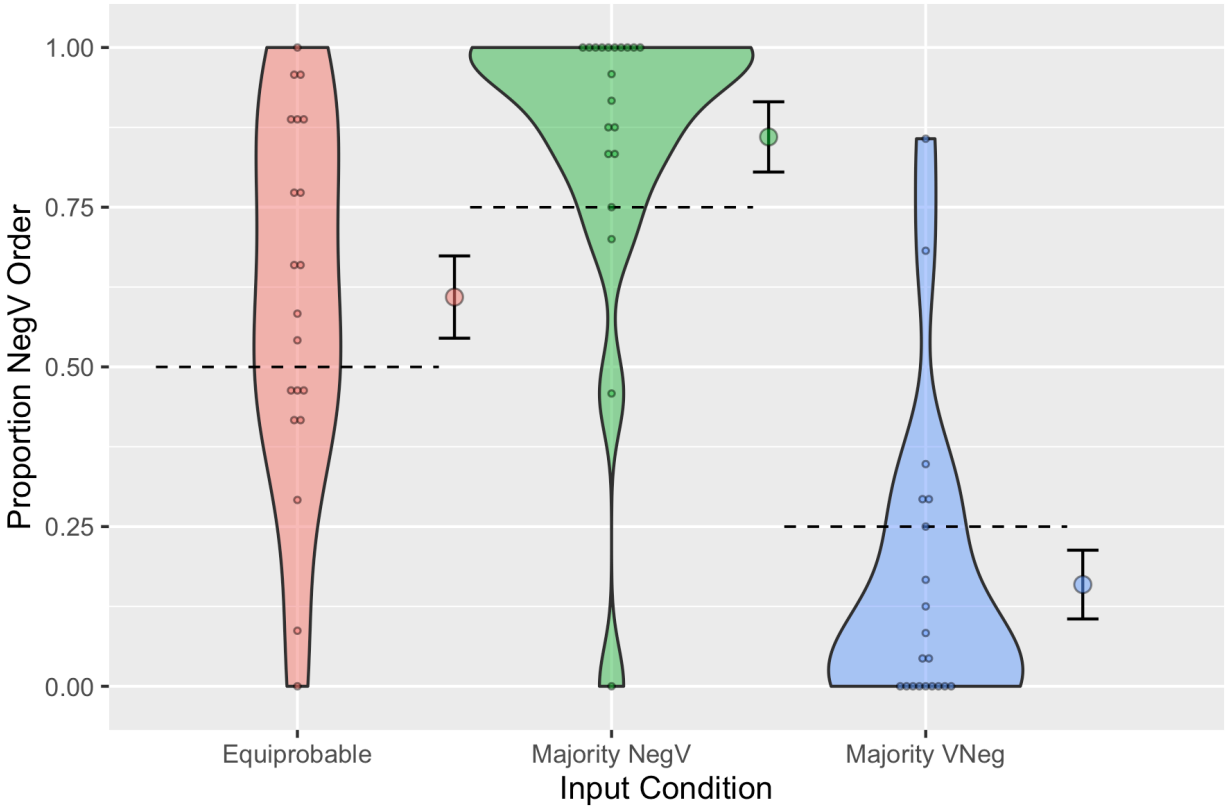
sentence responses which were greater than 4 edit distance away from a correct response was labeled as inaccurate and excluded from analysis; responses where there was not a clear closest match to either target sentence were also excluded from analysis.

Accuracy. Of the 67 participants that completed the experiment, recall that 7 were excluded because they did not meet a threshold of 75% accuracy in the final production phase. Among the remaining 60 participants, the mean accuracy was 95.5% in the final production phase; of the critical negation trials, 65 out of 1440 trials across all participants were excluded from analysis.

Across conditions and phases, overall accuracy in the forced choice task and production tasks was relatively high (95.4% in the forced choice task and 93.9% in the production task). However, to check whether there were differences in accuracy in either the forced choice or production tasks in different input frequencies condition, the buildmer package in R was used to evaluate the maximal feasible model for each of forced choice task accuracy and production accuracy, and to subsequently identify the model with the best fit through automated backwards stepwise elimination (Voeten 2022). This process compares the fit of a more complex model to the fit of a nested simpler model, using the Likelihood Ratio Test to determine whether the effect of a covariate is significant. The maximal effect structure for both the forced choice task accuracy and production accuracy models included Phase, Condition, and their interaction as fixed effects. For each, backwards stepwise elimination identified the model with only a fixed effect for Phase as the best fit. Notably, including Condition as a fixed effect did not result in a better fit, indicating that the input condition was not an important factor in explaining accuracy in either task. More detailed accuracy information across conditions and phases is available in Appendix D.

Neg-First Bias. The average proportion of productions with preverbal negation by condition are shown in Figure 9.

Figure 9. Proportion production of preverbal negation order by condition in Experiment 3. Dashed lines indicate proportion of preverbal negation order in the input; small points show subject means; the large points offset to the right of each violin plot indicate the overall mean for each condition; error bars indicate standard error of participant means.



To test for production preferences consistent with a Neg-First bias, negative sentence production data were analyzed using mixed effects logistic regression implemented with the lme4 package in R (Bates et al. 2015). To examine whether the proportion of preverbal negation productions in the Equiprobable Condition was significantly different than the input proportion of 50%, I ran a model with NegV word order use as the binary dependent variable, Condition as a fixed effect (reference level = Equiprobable), and random intercepts for Subject. This model produced a non-significant intercept ($\beta = 0.65$, $z = 1.2$, $p = 0.21$), indicating that the probability of productions with NegV word order was not significantly different from chance in this

condition. To test the prediction that participants in the Majority NegV Condition boosted the majority order significantly more than those in the Majority VNeg Condition, another model was run with Majority Order word order use as the binary dependent variable, Condition as a fixed effect (reference level = Majority NegV), and random intercepts for Subject and Item. This model found that the difference in the use of majority word order between the Majority NegV and Majority VNeg conditions was not significant ($\beta = -0.20$, $z = -0.25$, $p = 0.80$).

5.2.3 Interim Summary & Discussion

The results of Experiment 3 do not demonstrate the existence of a bias consistent with Neg-First preferences in adult Japanese speakers learning and producing a language which contained probabilistic variation of both preverbal and postverbal negation.

Notably, the results of Experiment 3 contrast with the biases found in L2 acquisition of Korean by Kim & Yun (2013), in which Japanese L1 learners of Korean produced more of the preverbal negation form than adults who spoke Korean as an L1, when tested with a written production task. Kim & Yun (2013) do not share the overall percentage of preverbal and postverbal negation constructions produced by each language experience groups in their study, but do report the average percentage of the postverbal negation construction used in each group across four different sentence conditions, showing that L1 Korean speakers produced the postverbal negation construction between 75% and 84% of the time across various sentence conditions, while L1 Japanese learners produced the postverbal negation construction between 17% and 38% of the time. There are several differences which could explain these diverging results:

1. We do not know the details of what and how the L1 Japanese learners of Korean were taught, except for the fact that they received instruction on both Korean negation

constructions. It could be that the learners in Kim & Yun's study received greater instructional emphasis on the preverbal negation construction than the postverbal negation one.

2. The preference of L1 Japanese learners to use the Korean preverbal negation construction could be driven by a preference to use particle negation, which can be interpreted as simpler: to transform the affirmative sentence to a negative one only requires the insertion of the particle to the left of the verb and does not require the conjugation of a different auxiliary verb, like the postverbal negation form. The artificial language of Experiment 3 eliminated this difference by using particle negation for both the preverbal and postverbal negation constructions.
3. There are several other differences between Korean (or any natural language) and the artificial language which could play a role in the divergent outcomes regarding preverbal negation production in Japanese-speaking learners of these languages. In the artificial language, the vocabulary size is very small, the exposure period brief, and the semantic function of negation limited to explicit denial of a visually presented transitive event. More empirical evidence from both natural language learning and artificial language learning studies, including more information linking how vocabulary size, length of exposure, and the presentation of grammatical elements affect learner behavior, would be useful to further understanding and aligning the mechanisms at play in each learning environment. I will elaborate on such future directions in Chapter 6.

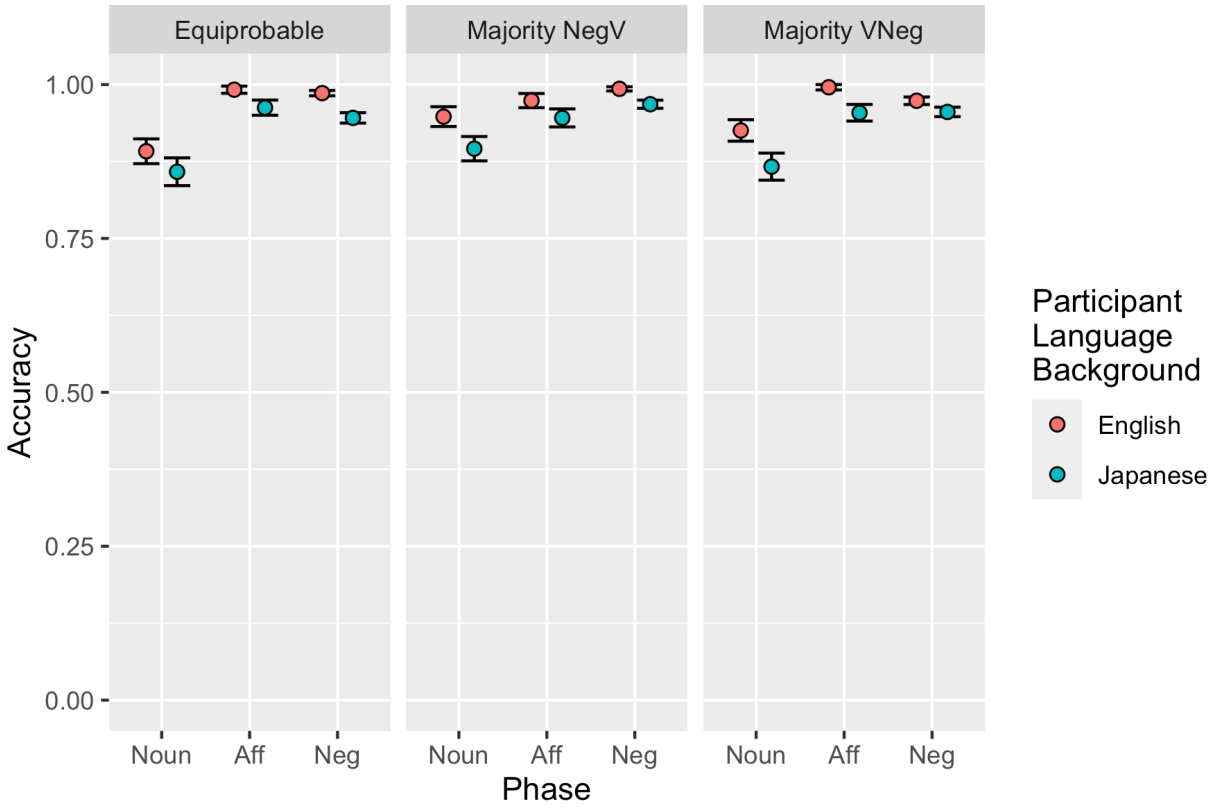
Next, I will turn to exploratory analyses comparing the results of Experiment 3 to Experiment 2 and examining the nature and effects of the Japanese-speaking participants' previous language experience (with an emphasis on experience with English).

5.2.4 Exploratory analysis: comparison to Experiment 2

As was previously noted, Experiment 3 is a partial replication of Experiment 2; though Experiment 2 contained a dyadic interaction task following the training and initial solo production task, the artificial language design and training procedures were nearly identical prior to dyadic interaction. Therefore, I combined the data from these experiments to compare how behavior differed between English and Japanese participant groups. Only data from the pre-interaction stages of Experiment 2 were included in these analyses.

Accuracy. Automated backwards stepwise elimination of logistic mixed effects regression models using the buildmer package for R was again used to confirm whether there were differences in accuracy performance in both the forced choice task and production tasks. For the forced choice task accuracy model, backwards stepwise elimination identified the model with only a fixed effect for Phase as the best fit, with fixed effects for Condition and Language Background not contributing to a better fit, indicating that the input condition and the participant's language background was not an important factor in explaining accuracy in the Forced Choice Task. However, for the production task accuracy model, the best fit model contained fixed effects for both Phase and Language Background. This model showed that the Japanese language background group had lower production accuracy overall than the English language background group ($\beta = -0.99$, $z = -3.66$, $p < 0.001$). Visual inspection of production accuracy results shown in Figure 10 shows a small but consistent and noticeable production accuracy difference between the different Language Background (i.e., Experiment) groups.

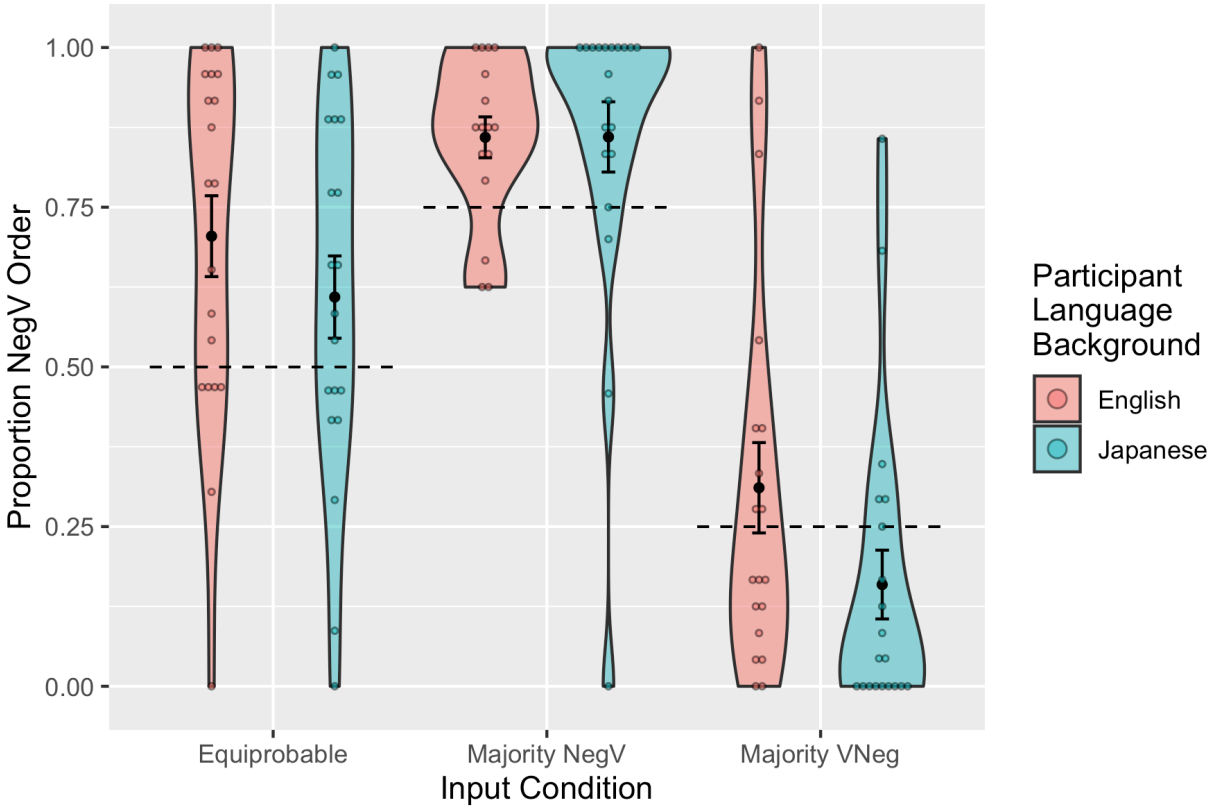
Figure 10. Production task accuracy production in Experiment 2 (English) and Experiment 3 (Japanese). Error bars indicate standard error.



Neg-First Bias. Figure 11 shows a comparison of the proportion of preverbal negation productions by Japanese speakers in Experiment 3 and English speakers in the first solo production stage of Experiment 2. To calculate whether there were pairwise differences between each of the two experiments within each condition, a mixed effects logistic regression model with NegV word order use as the binary dependent variable, fixed effects of Language Background, Condition and their interaction, and random intercepts for Subject. Pairwise comparisons were conducted on this model with Tukey-method p-value adjustment using the *emmeans* package (Lenth, 2020). These comparisons reveal that Japanese-speaking participants produced significantly more postverbal negation than English-speaking participants in the Majority VNeg Condition ($\beta = 1.69$, SE = 0.72, $p = 0.019$); there were no significant differences

between the two language background groups in the other input conditions (Equiprobable: β = 0.76, SE = 0.67, p = 0.25; Majority NegV: β = –0.58, SE = 0.77, p = 0.44).

Figure 11. Proportion of preverbal negation production in Experiment 2 (English) and Experiment 3 (Japanese). Dashed lines indicate proportion of preverbal negation order in the input; the transparent points show subject means; the black points indicate the overall mean for each condition; the error bars indicate standard error.



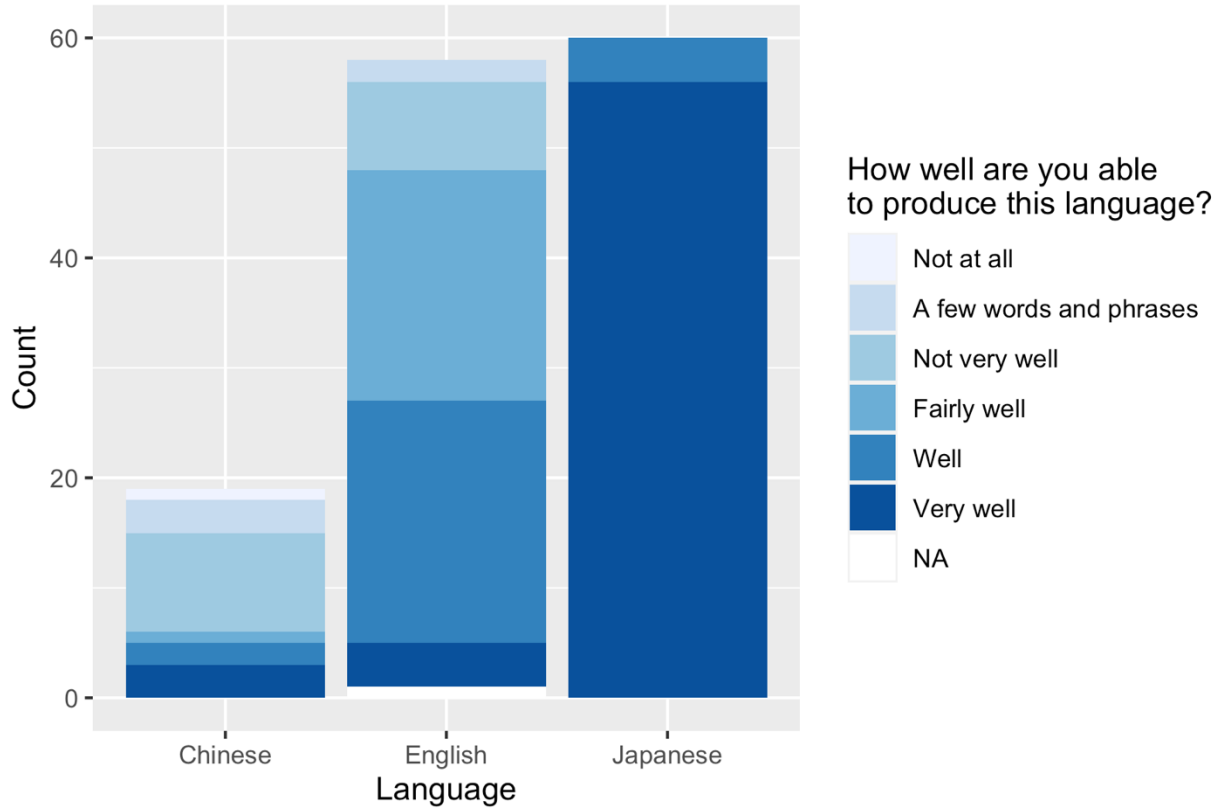
5.2.5 Exploratory analysis: experience with other languages

Given that the Japanese participants did not show a strong overall preference for postverbal negation, even though Japanese uses postverbal negation, there is reason to suspect whether experience with other languages, including English, may play a role in the pattern of results. Though all participants lived in Japan and identified themselves as having grown up speaking Japanese, some had considerable experience with other languages. Some prior experience with English was expected, as English is a common language of study in school and a

subject on university entrance exams in Japan (Glasgow & Paller 2016). Moreover, some theories of L3 language learning predict the L2 to be privileged as a source of transfer, over the L1, based on the assumption that L1 and L2 differ in the types of memory used (procedural for L1 and declarative for L2 and subsequent languages) (Bardel & Falk 2012). In this section, I explore whether participants with greater English experience had a greater preference for using NegV order in the context of the experiment.

Information about participants' language experience was collected through a post-experimental questionnaire. The questionnaire asked participants to name up to five significant languages (including Japanese) that they used or studied currently or in the past, and presented questions about age of acquisition, fluency in using and understanding the language, and frequency and types of exposure to the language (see Appendix B for list of questionnaire questions). In this section of the questionnaire, participants identified having some experience with English and (Mandarin) Chinese most often, outside of Japanese. Out of 60 participants, 58 named English and 19 named Chinese. To capture the general range of fluency participants had in these languages, Figure 10 shows the ratings participants provided for their ability to produce Japanese, English, and Chinese, if they named these as languages they had used or studied. Self-reported production ability, shown in Figure 12, was very highly correlated with self-reported comprehension abilities ($r(134) = 0.90$, $p < .001$) and moderately correlated with frequency of exposure through media ($r(132) = 0.70$, $p < .001$) and in social settings ($r(133) = 0.73$, $p < .001$).

Figure 12. Experiment 3 self-reported ratings of ability to "speak or produce" the top 3 languages that participants reported using or studying (Chinese, English, and Japanese).



Because English was a frequently named alternate language among participants, I examined whether self-rated proficiency in the ability to speak or produce English predicted participants' preference for NegV order. To do this, I ran a mixed effects logistic regression model with NegV use as the binary dependent variable, Self-Rated English Use Score, Condition, and their interaction as fixed effects, and random intercepts for Subject. A model including all 6 levels of self-rated English use scores failed to converge, so these scores were binned into two categories: all responses of 'well' and 'very well' were binned into one category representing 'higher' scores, while all remaining scores of 'fairly well' and lower were binned into another category representing 'lower' scores. Participants with lower self-reported scores were not found to differ significantly in their amount of NegV use from those with higher self-reported scores ($\beta = -1.86$, SE = 1.21, $p = 0.12$). Comparison of models with and without the

inclusion of Self-Rated English Use Score as a factor using the buildmer package showed that the inclusion of this factor did not improve the fit of the model.

5.2.6 Discussion

Unlike the English-speaking participants of Experiment 2, the Japanese-speaking participants of Experiment 3 did not show evidence of a bias favoring the production of preverbal negation over postverbal negation. This result contrasts with findings that Japanese-speaking learners of Korean use the option of preverbal negation in Korean more than the postverbal negation construction, and more than adults who speak Korean as a first language (Kim & Yun 2013). One possibility of several discussed in §5.2.3 is that the preference for preverbal negation in the acquisition of Korean is driven by differences in difficulty related to morphosyntax, which lead learners to prefer the preverbal negation construction, rather than any preference specifically to do with the order of the negator relative to the verb. Though there is still much to uncover about what enables or motivates a bias to produce preverbal negation in terms of how linguistic features of the target language interact with the learning context, the results of Experiment 3 do challenge arguments that tendencies for preverbal negation in creole languages, and language typology more broadly, are driven by a universal preference for preverbal negation in language acquisition (e.g., Bickerton 1981, Plag 2008, Schneider 2000). Instead, these tendencies may reflect properties of the languages and language experiences of the people involved in the histories of language contact and formation.

However, it is also interesting that the results of Experiment 3 also do not show evidence among Japanese-speaking participants of a bias favoring the production of postverbal negation; if we attribute the bias towards preverbal negation in Experiment 2 to participants' English language experience, we might also expect a bias towards postverbal negation in Experiment 3,

since participants had extensive experience with Japanese. There are some key differences between the two experiments which may factor into this result.

We find no evidence that that level of self-rated proficiency in English correlates with producing more of the English-like NegV order among Experiment 3 participants. However, it could be the case that any substantial experience with English increases the likelihood of producing NegV order in learning. Future studies with more rigorous comparison of populations with and without experience with language as an L2 would approach this question more directly (c.f. Culbertson et al. 2020).

The findings of Experiment 3 do not exclude the possibility that there is a universal bias for preverbal negation that is stronger in childhood but is diminished with long-term exposure to a postverbal negation language. Research on L1 acquisition of negation shows tendencies for children to overuse preverbal negation or learn preverbal negation before other negation constructions (though note the presence of preverbal preferences rely on some degree of variability in negation ordering in the target language, as in adult L2 acquisition; see §2.1.3 for discussion). Previous experimental work has found differences in adult and child language learning behavior have been found as the result of different cognitive capacities and linguistic knowledge. For example, work using artificial language learning paradigms has shown that children are more likely than adults to regularize inconsistent patterns of the linguistic input in production (Hudson Kam & Newport 2005, 2009; Schwab et al. 2018). Specifically, limitations on memory retrieval, which is more likely to affect children, has been suggested as one mechanism which leads to increased regularization (Hudson Kam & Chang 2009, Keogh et al. 2022). Future research incorporating children would be necessary to discern whether any

qualitative or quantitative differences affecting preferences in negation order exist between children and adults.

Though the initial training and production procedures in the artificial language in Experiment 2 and Experiment 3 were matched, Experiment 2 involved subsequent production tasks that Experiment 3 did not. Specifically, after an initial solo production task, participants in Experiment 2 went on to a dyadic production and comprehension task, in which they exchanged messages with another participant, in order to test a hypothesis that a preference for preverbal negation was driven by specific pressures of communication, and therefore would be more evident in contexts with an interlocutor. Even though the language training and solo production procedures were the same, an alternative explanation for the differences between the English-speaking participants in Experiment 2 and the Japanese-speaking participants in Experiment 3 is that the English-speaking participants produced more preverbal negation in anticipation of the dyadic production task, which they *were* aware would be coming up. I believe this is unlikely explanation for the source of this difference, since it remains unclear why there should only be a significant difference between experiments in the amount of preverbal negation produced within the Majority VNeg condition, but a replication with English speakers without the dyadic interaction task could test whether the presence of such a task affected the amount of preverbal negation produced in the previous solo production task.

Another way in which outcomes of Experiment 2 and 3 differed was in production error-rate. Though both Experiment 2 and Experiment 3 used similar typo-correction procedures and exclusion criteria, the Japanese-speaking participants of Experiment 3 generated significantly more production errors involving word labels being dropped or swapped. Though efforts were made to match difficulty and familiarity across the two experiments by converting the artificial

language to Japanese script and phonotactics, the increased error rates signal that the resulting language may have still been more difficult or perceived as more unusual by Japanese speakers than the corresponding language was for English speakers, as previous artificial language learning studies have found evidence that less familiar form-function mappings negatively affect lexical accuracy (Baptista et al. 2014). However, it should be noted that while the production error difference was significant, it was relatively small, and there was no difference in accuracy in the 2-Alternative Forced Choice task intended to test comprehension. While I would be interested in further investigating in future work the effect of phonological, syntactic, and orthographic similarity (based on perceptions of the learner or formal dimensions laid out by linguists) on difficulty and its interactions with other biases in learning and production, such as the regularization bias, I think the subtle difference in error rates is unlikely to play a large role in explaining the overall pattern of results seen here.

5.3 Conclusion

Though a bias to produce preverbal negation in the early stages of language acquisition has long been presumed to be universal, Experiment 3 provides evidence that previous language experience modulates this bias in adult learners. Whereas Experiment 2 demonstrated a bias in favor of preverbal negation among English-speaking learners, this bias was not found among Japanese-speaking learners in Experiment 3. Given the lack of a reverse tendency among participants to prefer postverbal negation, it is worth noting that these findings could be compatible with the idea that a Neg-First bias may be present in childhood but diminish with experience with a postverbal negation language like Japanese. Conversely, it could be the case that a postverbal negation bias resulting from Japanese experience was mitigated by having any experience with preverbal negation languages like English, given that L2 knowledge may be a

privileged source of language transfer (Bardel & Falk 2012). However, at face value, the present findings begin to call into question whether a universal bias in acquisition can explain the prevalence of preverbal negation in the world's languages, even in genetically and areally balanced samples (Dryer 1988). Instead, it may be fruitful to look to alternative or more nuanced explanations for the cross-linguistic prevalence of preverbal negation. It may emerge from historical forces, unrelated to cognitive preferences related to ease of acquisition or use, or it may emerge from separate cognitive preferences such that *apparent* preferences for preverbal negation are actually rooted in preferences for other linguistic elements (e.g., particle negation) or appear only in conjunction with other linguistic features which commonly appear in constellation with preverbal negation (e.g, SVO constituent order).

Chapter 6 Summary, Implications, and Future Directions

The tendency for negation to appear early in the sentence, dubbed the "Neg-First principle" by Horn (1989:452), has been observed in the domains of typology, language contact, and language acquisition. While scholars have speculated about the source and universality of the Neg-First principle based on such evidence, these hypotheses have not been widely tested using experimental methods. In this dissertation, I integrated and assessed the contributions and limitations of the research on negative word order in these various disciplines, and used artificial language learning to experimentally investigate how these findings might be explained by biases at play in language learning and communication. These experiments sought to address whether there was evidence of a bias to produce early negation, whether a communicative context encouraged this bias, and the degree to which this bias was shaped by previous language experience. Here I summarize the key research findings and the contributions thereof and propose avenues for future research.

6.1 Considering Possible Sources of Neg-First Tendencies

In Chapter 2, I compiled varied evidence for the Neg-First principle from the fields of typology, language contact, language acquisition, and psycholinguistics, as well as explanations and mechanisms proposed to account for the Neg-First principle and other aspects of negation order. Mechanisms proposed to affect negation order included universal constraints on processing and production (e.g., Processability Theory, Pienemann 2008; Production-Distribution-Comprehension account, MacDonald 2013), principles of ambiguity avoidance and audience design in communication (e.g., Dryer 1988, Horn 1989), and grammaticalization (e.g.,

Jespersen's Cycle). I also described some of the challenges and limitations of relying on only typological and acquisition evidence to infer a universal Neg-First principle and its underlying mechanisms. Observations of typological distributions and patterns are subject to uncertainty about whether typologically common patterns are more common due to the advantages they may afford in cognition or use, or due to potential non-independence of the sample and historical factors of language inheritance. In acquisition, since languages hardly ever only differ as to whether the negator appears before or after the verb, it is difficult to confirm whether a preference to produce preverbal negation reflects a preference for preverbal negation *per se* instead of a preference for other features which factor into the choice to use preverbal negation. Furthermore, it is difficult or impossible to know the exact nature of the learners' input, and therefore to learn how the learners' biases differ from or rely on the frequencies and properties of their input. Therefore, I introduce three artificial language learning experiments to test for the presence of Neg-First tendencies in acquisition. Though these experiments differed as to what they required from participants, they were used to investigate the possible sources of Neg-First tendencies in typology across different populations, while allowing for features of the language and learner's input to be carefully controlled.

Through these three experiments, I tested the presence of Neg-First biases in the early stages of second language acquisition by examining whether participants exposed to an artificial language containing both preverbal and postverbal negation demonstrate a bias to overproduce preverbal negation compared to their training input. Experiment 1, testing whether English-speaking participants demonstrate a preference consistent with a Neg-First bias, produced inconclusive results. Participants did show a small numerical tendency to produce more

preverbal negation than was present in the training input, but this effect was not statistically significant.

Experiment 2 addressed potential limitations of Experiment 1 by adding a dyadic interaction component, to test the hypothesis that Neg-First biases are driven by a communicative desire to reduce potential for the listener to misunderstand the intention of the sentence, and by removing the presence of a word bank to encourage greater regularization by making lexical retrieval more difficult. The results of Experiment 2 did indicate a bias to overproduce preverbal negation among English speakers, consistent with a Neg-First bias. However, interaction did not boost the use of preverbal negation relative to the solo production phase, except in the Majority Preverbal Negation Condition, in which the bias to regularize the majority order in the training language was consistent with a bias to use preverbal negation. This result appears consistent with prior research indicating that mechanisms of structural priming and asymmetric accommodation lead dyads to converge on regular systems, particularly in the presence of a strong bias towards a specific pattern, but it is not consistent with the idea that a communicative context induces a greater preference for earlier negation in general.

Finally, Experiment 3, a partial replication of Experiment 2, sought to investigate the role of previous language experience on the preference to produce preverbal negation in adult L2 acquisition, by testing whether speakers of a postverbal negation language (Japanese) showed a preference to produce preverbal negation in the early stages of learning. The results showed that a Neg-First bias was modulated by language experience: while the English-speaking participants of Experiment 2 showed a preference to produce preverbal negation, Japanese speakers did not. As I will elaborate in the following section, this challenges the idea of a universal preference for preverbal negation in acquisition as a plausible candidate for shaping linguistic typology.

Though one possibility is that the cross-linguistic prevalence of preverbal negation results from accidental consequences of language history rather than pressures on learning and use, more nuanced explanations which maintain the role of biases present independent of language experience remain viable avenues for future exploration.

6.2 Implications for Understanding Language Change and Variation

This dissertation contributes to a larger body of research which aims to motivate and test inferences and theories based on typological distribution with patterns of behavior in individuals using artificial language learning and demonstrates the value of doing such work with varied populations. Though typologists and creolists have sometimes speculated about the existence of a universal bias for preverbal negation in language acquisition or language processing in order to explain word order tendencies, the results of Experiment 3 caution against this conclusion. In particular, the finding that English speakers, but not Japanese speakers, show evidence of a bias favoring the production of preverbal negation after being trained on a language that contained both pre- and postverbal negation in free variation challenges the idea that there is a universal (or at least language non-specific) bias in adult L2 acquisition driving the typological tendency for languages to use more preverbal negation. Instead, the typological tendency to use preverbal negation may reflect properties of the languages and language experiences of the people involved in the histories of language contact and formation. This reflects the challenge of understanding how different language experience interacts with properties of the target language and general mechanisms of language acquisition and use to drive language change and variation.

Let us also consider that grammaticalization can be a source for word order correlations that would not necessarily be represented in synchronic universals of language acquisition or language processing; rather word order correlations derived by grammaticalization processes

reflect the word order correlations of the historical source of an element (e.g., negators are often derived from quantifiers) which themselves might be correlated due to language-internal historical reasons or due to constraints on learning or processing. However, as described in §2.2.3, grammaticalization pathways do not neatly explain the cross-linguistic prevalence of preverbal negation in VO languages, since typical sources of negation would be found in object position, after the verb.

There are still alternative explanations that could be raised for apparent tendencies to use preverbal negation in typology and acquisition, some of which could still retain the role of universal biases at play in language processing and acquisition. The fact that Japanese speakers in Experiment 3 did not show an inverse preference for postverbal negation could be consistent with a Neg-First bias that is present at childhood but diminished by experience with a postverbal negation language, or a postverbal negation preference based on Japanese experience that is diminished by experience with English in school and media, even if they did not rate themselves as proficient users of English at the time of the study. Though such avenues could be explored in future research, all of these possibilities underscore the importance of examining the role of language experience, development, and contact together when investigating the role of universal principles in propelling language variation and change.

Artificial language learning is not only a useful tool to establish whether there are behavioral biases in learning, and how these appear in different populations, but experimental data and the ability to manipulate properties of the learner's input and the learning environment is important for understanding the mechanisms at play in language learning and contact. The introduction of a dyadic director-matcher paradigm in Experiment 2 contributes to existing scholarship examining how regularization is affected by mechanisms at play in interaction, such

as priming and (asymmetric) accommodation, and provides empirical data showing how these mechanisms play out in the presence of a substantial preference for a certain variant. However, Experiment 2 did not find evidence consistent with an audience design explanation for a Neg-First bias. Given this, we might consider how this bias might arise from mechanisms which center producer-related difficulty and efficiency, rather than mechanisms which center comprehender difficulty. Transfer from the L1 and the ability to coordinate word order between the L1 and the L2 would provide a language-experience-dependent, producer-centered explanation for Neg-First biases when the L1 has preverbal negation. If future work determines that Neg-First biases do go above and beyond L1 transfer explanations, the unique contextual or conceptual salience of negation might be another path for explaining a Neg-First bias in production.

Finally, this thesis provides a model for motivating and testing a purported cognitive bias through the application and evaluation of evidence from typology, acquisition, language contact, and artificial language learning. I hope that I have demonstrated in this dissertation the importance and utility of carefully collecting and evaluating converging, compatible, and contradictory evidence across different domains and methodologies in the pursuit of understanding the contributions of various processes, such as language acquisition and use, in language change which shape language patterns over time.

6.3 Avenues for Future Research

6.3.1 Conditioned Variation and Neg-First Biases in Language Learning

A prominent finding in the negation acquisition literature is that children and adults tend to show developmental patterns in which preverbal negation is erroneously favored in their productions only when there is variation in the order of negation in the target language.

However, variation in natural language is often *conditioned*, so that variation is rarely fully unpredictable, but the variation in the experiments presented in this dissertation were not conditioned on any other linguistic features present in the training (e.g., NegV order did not occur more frequently with a certain set of verbs). Variation in negation order (relative to the main verb) in natural language may be conditioned on any number of linguistic, pragmatic, or sociolinguistic features. For example, in V2 languages like German, the order of negation relative to the main verb depends on clause type and whether an auxiliary verb is present. In main clauses, the negator appears before the main verb when an auxiliary verb is present (32a), but after the main verb when no auxiliary verb is present (32b). In subordinate clauses, the negator appears before all verbs in the clause (32c).

(32) a. Laura hat **nicht** geschlafen.
Laura has-3SG NEG slept
'Laura didn't sleep.'
b. *Laura schläft* ***nicht****.*
Laura sleep-3SG NEG
'Laura doesn't sleep.'
c. *Laura ist müde, weil sie* ***nicht*** *geschlafen hat.*
Laura COP-3SG tired because 3SG NEG slept has-3SG
'Laura is tired, because she didn't sleep.'

Negation placement conditioned on linguistic and social factors such as clause type or register is relevant to patterns of obligatorification and loss in Jespersen's Cycle and grammaticalization more broadly. For example, Welsh has undergone Jespersen's Cycle, with different present-day results in in different clause types. Present-day spoken Welsh expresses standard negation in with the marker *ddim* in (28a), which follows both the verb and subject, and originated as a negative polarity item ('at all'). *Ddim* has entirely supplanted an original, preverbal negative marker *ni(d)* in spoken Welsh main clauses (28b). However, the original preverbal negator is generally retained in subordinate clauses, and *ddim* is not required in these

subordinate clauses. Thus, in spoken Welsh, Jespersen's Cycle has progressed further in main clauses than in subordinate clauses: *ddim* has supplanted the original negator *ni(d)* (Middle Welsh *ny(t)*) in main clauses, but in subordinate clauses the original negator is retained.

(33) Welsh (Willis 2010: 111)

a. *Ddywedodd y gŵr* ***ddim*** *wrth Peredur beth oedd hynny.*
say.PST.3SG the man NEG to Peredur was be.IMPF.3SG this
'The man didn't tell Peredur what this was.'

Middle Welsh (Peredur 20.11–12, in Willis 2010: 111)

b. ***Ny*** *dywawt y gwr y Peredur beth oed hynny.*
NEG say.PST.3SG the man to Peredur what be.IMPF.3SG this
'The man didn't tell Peredur what this was.'

To demonstrate further examples of the types of conditioned variation which affect negation placement in natural languages, recall that in Korean, corpus studies have demonstrated that the choice between the preverbal and postverbal negation constructions shows some degree of conditioning based on register (preverbal negation is associated more with spoken, colloquial language) and lexical item (preverbal negation is less frequently applied to Sino-Korean predicates and adjectival verbs with at least three syllables).

Learners acquiring languages with such variation need to learn the appropriate conditioning environment to consistently order negation in a target-like manner. When this is particularly difficult for learners to do, perhaps because the features conditioning negation are, for whatever reason, not salient to the learner, then we may anticipate that underlying biases may play a greater role in learners' variant choice.

Artificial language learning has already been used to look at the learning of conditioned variation. For example, previous studies have shown that both children and adults could successfully learn lexically conditioned variation (e.g., Hudson Kam & Newport 2009, Wonnacott 2011). Samara et al. (2017) found that both adults and children have also been shown

to be sensitive to variation conditioned on speaker identity, even when this conditioning was probabilistic and only partially consistent. Moreover, Samara et al. (2017) examined how participants learning a language *introduced* conditioning in different learning contexts, finding that adults, but not children, tended to introduce lexically-conditioned variation which was not present in the input, whereas neither group tended to introduce speaker-conditioned variation which was not present in the input. Brown et al. (2021) examined learning and generalization of grammatical regularities that correlate with semantics, exploring how well six-year-old children and adults learned particle usage that is fully or partially determined by noun semantic class. Brown et al. (2021) found that both groups generalized to novel nouns when semantic cues were fully consistent, but children failed to generalize when cues were only partially consistent. Work in reinforcement and error-driven learning which explores how learners attend to and weigh different cues in second language acquisition (e.g., Harmon et al. 2019, learning which phonetic cues reliably indicate phonemic category distinctions) can also be applied in this domain to understand how learners track and model the conditioning environments of syntactic variation.

Research from artificial language learning and applied linguistics which explores how well different types of conditioned variation are learned in different linguistic and social contexts and how this interacts with patterns of regularization and generalization, would aid in developing a more complete model of how and when Neg-First biases (and other biases) emerge in acquisition. Specifically, looking at the learning of deterministically- and probabilistically-conditioned variation of negation placement when conditioning depends on register (formal or informal, spoken or written), clause type (main or subordinate), and lexical item (especially verbs or verb classes) would have direct connections to the types of variation that are relevant to identify and produce in natural language acquisition of negation. Looking at whether individuals

with different language backgrounds, for whom different types of conditioned variation may be more or less relevant in their previous language(s) would also be of interest.

6.3.2 Enabling Causes and Agents of Language Change

Kiparsky (1996) posited that "language internal" linguistic change involved an interaction of *efficient causes* (a preference for a certain linguistic feature or pattern) and *enabling causes* (something which caused the old pattern to be less entrenched or less attainable). This dissertation has been primarily concerned with investigating motivations for *efficient causes* for Neg-First biases; in the artificial language learning paradigm used, the *enabling cause* was already granted by presenting the two-word order variants of interest in free variation. Of course, variation in natural language is usually conditioned; however, changes to "the sociolinguistic context, the radius of communication, or the existence of diglossia or multiple norms" may make conditioning factors opaque to language users, enabling efficient causes to take hold (Kiparsky 1996: 14).

Whether we are concerned with explaining cross-linguistic or local trajectories of language change, future research should consider what social and linguistic conditions enable subtle pressures on language learning and use to enact change. In this vein, debates around *who* is positioned to affect broader community-level language change—specifically, whether children and adult L2 language learners can meaningfully affect language change is of crucial importance (e.g., Aboh 2015, Lupyan & Dale 2010, Dale & Lupyan 2012, Lightfoot 1997). Language change does not only involve learning, but also production and transmission, and it is the case that different types of speakers who may be more 'prototypical learners' may not also be 'prototypical transmitters'. For example, based on errors made by L2 learners of Swedish, Hyltenstam (1977) conjectures that the starting point of acquisition for all learners involves

placing the negator immediately before the finite verb in main clauses. However, none of the participants in Hyltenstam's study did so ubiquitously, even though the first test was performed after only three weeks of study. In fact, among learners who invariantly placed the negator on the same side of the verb regardless of verb type or clause type, they all placed negation after the finite verb. Though differentiating the main and subordinate clause environments creates difficulty affecting accurate placement of the negator, most learners will not stay at that point very long, with postverbal negation in main clauses as the norm. This issue has been raised to contest the idea that early language learning stages of L2 learners can meaningfully affect language change, even in emerging creole contexts (Aboh 2015).

Under the assumption that the language of young children is not readily adopted by adult speakers (note that errors in negation placement made by children are typically resolved before age 4) the relevance of children as agents of language change has also been contested (e.g., Arends & Bruyn 1995, Kerswill & Williams 2000). Nevertheless, independently of considerations for linguistic innovation and diffusion, studies examining the qualitative and quantitative differences in the biases between and among child and adult learners can shed light on subtle learning mechanisms and trajectories across the lifespan. For example, Culbertson et al. (2012) found that adults learning nominal phrases in which numeral and adjective modifiers were on different sides of the noun ('non-harmonic' orders) struggled to learn the typologically rare Adj-N, N-Num order, instead tending to shift from post-nominal to pre-nominal numerals. Because the participants of the study were English-speaking, Culbertson et al. (2012) suggested that this might be a reflex of participants' native language preference. However, taken together with later evidence from children's learning outcomes, we see that this could instead reflect a more general tendency for learners to generalize the adjective order to numerals rather than the

reverse (Culbertson & Newport 2015). Children generalized to the adjective order for both non-harmonic orders (e.g., children learning N-Adj, Num-N master N-Adj and generalize to N-Num; children learning Adj-N, N-Num master Adj-N and generalize to Num-N), whereas adult learners only showed this generalization for the rarer Adj-N, N-Num pattern. Taken together, these experiments reveal that a harmonic bias is markedly stronger in children than adults, though an intriguing tendency to generalize the adjective order may be common to both children and adults. Results such as this underscore the range of factors that may combine to affect the shape of language and the direction of language change. Whether Neg-First biases in acquisition are universal or language-specific, further, sophisticated work on how age structure affects learning and social diffusion is necessary to fully develop models linking individual-level biases to community-level language change (cf. Christiansen & Chater 2016, Roberts & Sneller 2020).

6.4 Conclusion

This dissertation presented initial steps towards testing proposed links between Neg-First tendencies in typology and biases in language learning and use using artificial language learning. Though English-speaking participants displayed a bias to produce preverbal negation when learning an artificial language with both preverbal and postverbal negation options in free variation, Japanese speakers did not clearly show this bias. Though this doesn't necessarily refute the possibility of a language-experience independent Neg-First bias in language learning or use that can explain the prevalence of preverbal negation in the world's languages, it does suggest that we should be cautious of that claim, and further investigate the impact of language experience on Neg-First biases. While the tendency for languages to place the negator preverbally has been proposed to derive from a functional preference to place the negative marker early in the sentence to facilitate the ability for the recipient to process the negative

sentence (Dryer 1988, Horn 1989), such an explanation may not be necessary. I leave it to future work to further consider the extent to which a behavioral preference for preverbal negation exists among populations with different language backgrounds and age cohorts. Future work should also further consider the relevance of other linguistic features which commonly appear in conjunction with preverbal negation (e.g., SVO constituent order, particle negation), and whether Neg-First preferences need to be motivated in terms of the specific semantic status of negation, or whether the observed tendencies can be motivated by historical forces and cognitive biases that apply more broadly.

Appendix A Experiment 1 Post-Experiment Questionnaire

1a. Did you grow up in the United States? Yes / No

1b. If you answered No to 3a, where did you grow up?

2. What languages did you speak/hear growing up?

3. If you have lived or travelled in countries other than the United States for three or more months, then indicate the name of the country, length of stay, the language you used, and the frequency of your use of the language for each country.

Country	Length of stay[a] [month(s)]	Language	Frequency of use[b]
			1 2 3 4 5 6 7
			1 2 3 4 5 6 7
			1 2 3 4 5 6 7
			1 2 3 4 5 6 7
			1 2 3 4 5 6 7

a. You may have been to the country on multiple occasions, each for a different length of time. Please add all the trips together.

b. Please rate according to the following scale: (circle the number in the table)

Never	*Rarely*	*Sometimes*	*Regularly*	*Often*	*Usually*	*Always*
1	2	3	4	5	6	7

4. Have you ever studied or learned a second language in terms of listening, speaking, reading, or writing? (Circle one): Yes / No

5. Indicate your native language(s) and any other languages you have studied or learned, the age at which you started using each language in terms of listening, speaking, reading, and writing, and the total number of years you have spent using each language.

Language	Listening	Speaking	Reading	Writing	Years of use[a]

a. You may have learned a language, stopped using it, and then started using it again. Please give the total number of years.

6. Is there anything else you would like to tell us about your experience with languages?

7. What do you think this study was about?

8. As you were trying to learn the language in this study, did you notice yourself using particular strategies to help you learn it? If so, can you describe them?

9. How easy or difficult did you find it to learn this language?

Appendix B Experiment 2 and 3 Post-Experiment Questionnaire

Upon completing Experiments 2, participants were automatically redirected to a post-experiment questionnaire hosted on Qualtrics. The Experiment 2 questionnaire was translated into Japanese for Experiment 3 by Dr. Yushi Sugimoto.

In this section, please reflect on the negative sentences you saw in training & your hypotheses while you were learning.
このセクションでは、トレーニングで見た否定文に関してあなたが思いついた仮説について聞いていきます。

1. How easy or difficult did you find it to learn this language?
   この言語の学習は、どのぐらい簡単、あるいは、難しかったですか？

2. While you were learning how to produce negative sentences in this experiment, do you remember using any particular strategies or considering certain rules? Did you come to any conclusions about how negation was formed?
   実験において，否定文の使い方をどのように学習したか，あるいは，どのようなルールがあると思ったかを思い出してください。否定がどのように使われているかについて何かわかったことがありますか？

3. What percentage of the time do you think “pik” appeared to the left of the verb (e.g., “pik umi mook daki”) in your training?
   トレーニングの中で「ピク」という単語がどのぐらいの頻度で動詞の左にきたと思いますか？（例 ピク トマ シド ホレト)

4. What percentage of the time do you think “pik” appeared to the left of the verb (e.g., “pik umi mook daki”) in your own production?
   産出(単語のアウトプット)した際「ピク」という単語がどのぐらいの頻度で動詞の左にきたと思いますか? (例 ピク トマ シド ホレト)

5. Did you finish the interactive game? [Note: not asked in Experiment 3]
   - Yes, I finished the interactive game.
   - No, I did not finish the interactive game.
   - No, I did not play the interactive game; I timed out on the pairing screen.

6. (If yes response to Q5) Did the interactive game change your hypotheses about how negation worked at all?

This section contains questions about your language history. If you would like to clarify any of your answers, you will have a chance to do so at the end of this section.
このセクションではあなた自身に関する質問とあなたが学んできた言語について聞きたいと思います」。もし自分の答えを変えたいときは、このセクションの最後に変更が可能となります。

7. Did you grow up in the United States?
   - Yes
   - No

   日本で育ちましたか？ [Did you grow up in Japan?]
   - はい
   - いいえ

8. (If no response to Q7) Where did you grow up?
   どこで生まれ育ちましたか？

9. What language(s) did you use while growing up, at home or in school?
   子供の時, どの言語を家あるいは学校で使っていましたか？

10. Indicate your first language and any other languages you have studied or learned, the **age at which you started** learning or using each language, and the total number of years you have spent using each language*. Your best guess is fine. If you were exposed to a language since birth, you may put age 0 for age you began learning.

    If you have studied or learned fewer than five languages please leave extra rows blank; if you have learned more than that number please choose the five languages that you use most frequently or that you think are most influential to you.

    *You may have learned a language, stopped using it, and then started using it again. Please give the total number of years of active use.

    "あなたの第一言語、 今まで学習した言語を記入してください。また学習した言語に関していつから学習を始めたか、 そして学習の期間も記入してください。 学習していない期間もある場合は、 学習した合計期間を記入してください。 推定の期間で構いません。

    日本語も学習した言語として記入してください。 生まれた際に他の言語の影響を受けた場合、 0 歳からその言語の学習が始まった期間として記入してください。

    学習した言語が 5 個より少ないの場合, 余白の行は残してください。 学習してきた言語が 5 個以上の場合、 もっとも学習してきた言語，あるいは，最も影響を受けた言語を 5 つ選んでください。"

	Language	Age you began learning	Years of use
1	English		
2			
3			
4			
5			

11. For each of the languages you listed above, please answer how well you are able to **speak or produce** that language now.
上記に挙げた言語それぞれに関して, 現在どのぐらい話す, あるいは使用することができるかを答えてください.

Not at all	No more than a few words or phrases	Not very well (I can only talk about simple/basic things)	Fairly well (I can talk about some things)	Well (I can talk about many things)	Very well (I can talk about almost anything)
○	○	○	○	○	○

12. For each of the languages you listed above, please answer how well you are able to **understand** that language now.
前のページで記入された言語それぞれに関して, どのぐらいその言語を理解できるかに関して答えてください.

Not at all	No more than a few words or phrases	Not very well (I can only talk about simple/basic things)	Fairly well (I can talk about some things)	Well (I can talk about many things)	Very well (I can talk about almost anything)
○	○	○	○	○	○

13. For each of the languages you listed above, how often do you think you are exposed to the language in your daily life, by means of radio, TV, or online media?
上記で記入された言語それぞれに関して, 日常生活でどのぐらい触れているか答えてください（例えば、ラジオ、テレビ、オンラインのメディア等）.

Less than once a year	Less than once a month	Less than once a week	Less than once a day	At least once a day
○	○	○	○	○

14. For each of the languages you listed above, how often do you think you are exposed to the language in conversation at work, at home, or in social settings?
上記で記入された言語それぞれに関して, 仕事場, 家, あるいは, 社会的な場面にてその言語にどのぐらい触れているか答えてください.

Less than once a year	Less than once a month	Less than once a week	Less than once a day	At least once a day
○	○	○	○	○

15. Is there anything else you would like to tell us about your experience with languages?
言語に関するあなたの経験について特筆すべきことがあれば記入してください。

Appendix C Experiment 2 Accuracy Scores

Figure C.1. Mean accuracy across forced choice task (comprehension) trials in Experiment 2. Small white dots show individual scores; black dots represent the mean within a given training experimental condition and phase; error bars represent standard error.

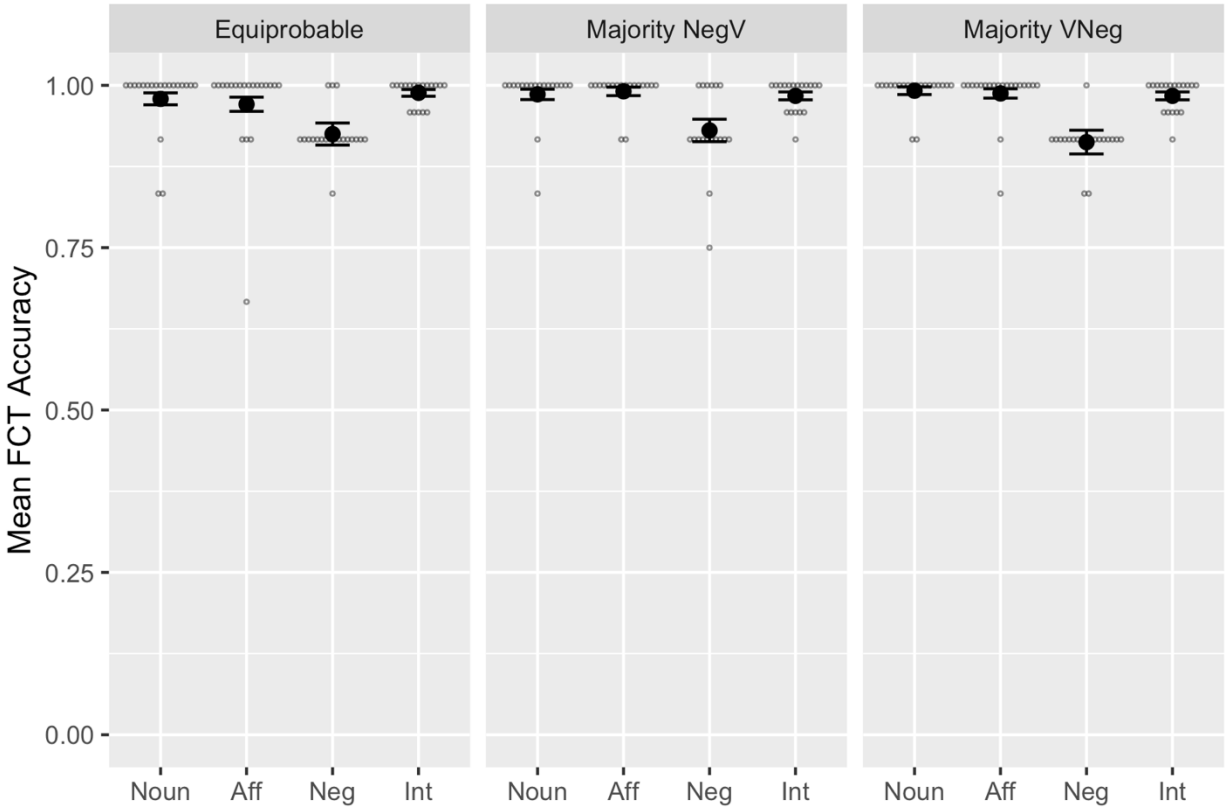
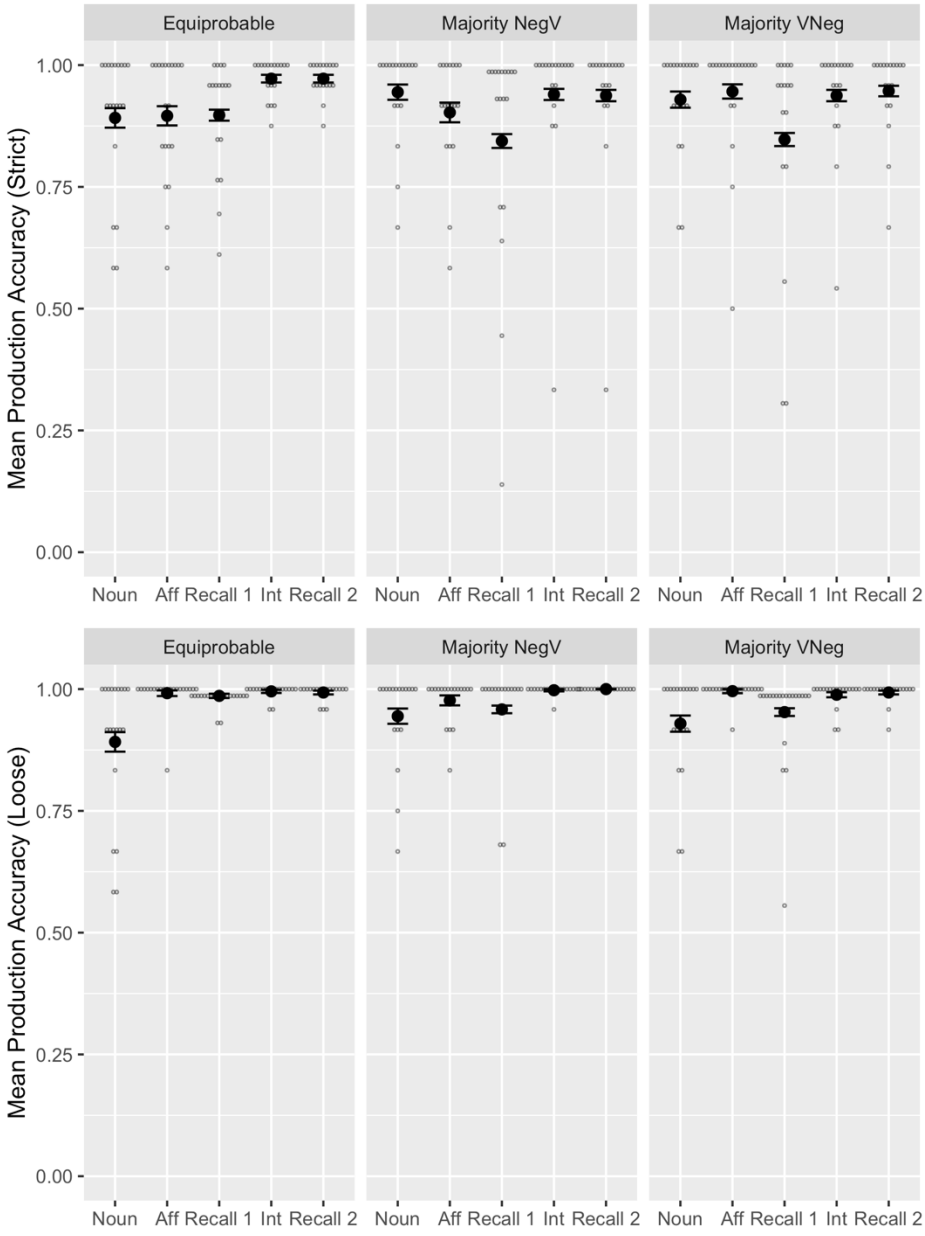


Figure C.2. Mean accuracy (strict and loose measures) across production trials in Experiment 2. Small white dots show individual scores; black dots represent the mean within a given training experimental condition and phase; error bars represent standard error.



Note: 'Strict' accuracy scores were calculated as follows: 1 for exact matches between the participant input and target, and 0 for all other responses. 'Loose' accuracy scores were calculated: 1 when the Levenshtein distance from the target string was 5 or less, and 0 for all other responses. For instance, a response where only one lexical item was swapped out for another would be classified as accurate by the loose accuracy measure. Both strict and loose accuracy measures were calculated after correcting for minor typos (i.e., eliminating punction, converting to lowercase, and converting words which were only 1 character different (using Levenshtein Distance) from a legal word to the legal word

Appendix D Experiment 3 Accuracy Scores

Figure D.1. Mean accuracy across forced choice task (comprehension) trials in Experiment 3. Small white dots show individual scores; black dots represent the mean within a given training experimental condition and phase; error bars represent standard error.

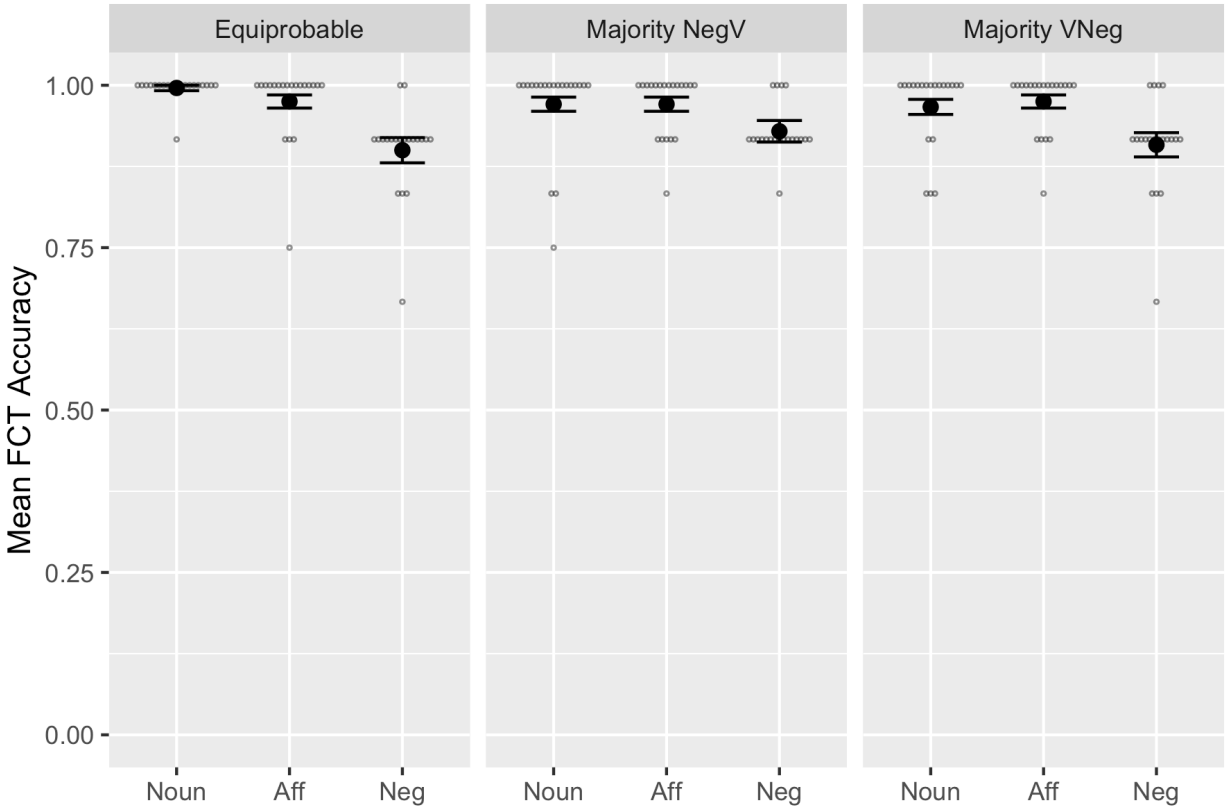
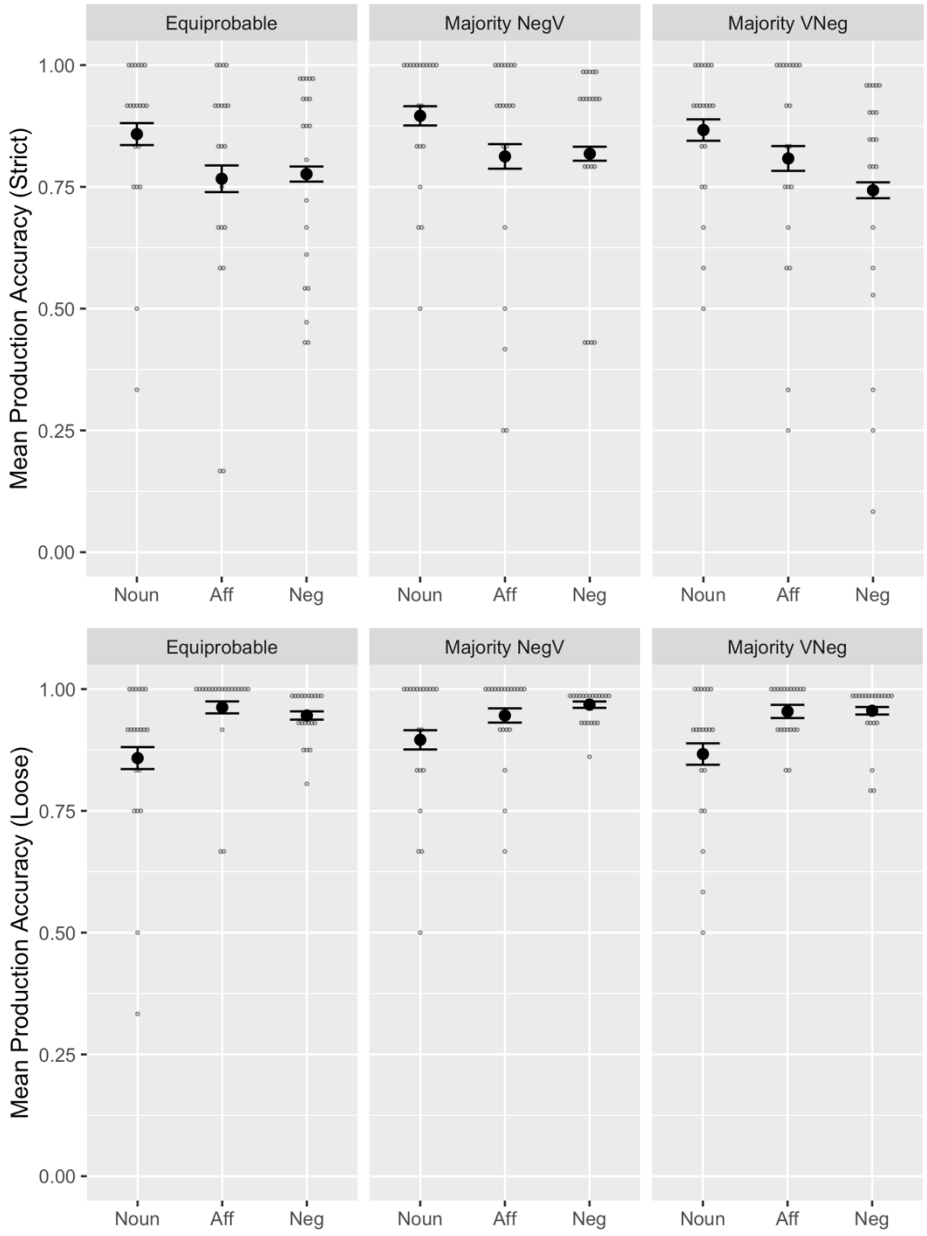


Figure D.2. Mean accuracy (strict and loose measures) across production trials in Experiment 3. Small white dots show individual scores; black dots represent the mean within a given training experimental condition and phase; error bars represent standard error.



Note: 'Strict' accuracy scores were calculated as follows: 1 for exact matches between the participant input and target, and 0 for all other responses. 'Loose' accuracy scores were calculated: 1 when the Levenshtein distance from the target string was 4 or less, and 0 for all other responses. For instance, a response where only one lexical item was swapped out for another would be classified as accurate by the loose accuracy measure. Both strict and loose accuracy measures were calculated after correcting for minor typos (i.e., eliminating punction, converting to lowercase, and converting words which were only 1 character different (using Levenshtein Distance) from a legal word to the legal word.

Bibliography

Aboh, Enoch. 2010. C-Type Negation Markers on the Right Edge. In Aboh, E., Essegbey, J. (eds.), *Topics in Kwa Syntax*. Studies in Natural Language and Linguistic Theory, vol 78. Springer, Dordrecht.

Aboh, Enoch. 2015. *The emergence of hybrid grammars: Language contact and change.* Cambridge: Cambridge University Press.

Aboh, Enoch & Michel DeGraff. 2017. A Null Theory of Creole Formation Based on Universal Grammar. *The Oxford Handbook of Universal Grammar*, 400–458. https://doi.org/10.1093/oxfordhb/9780199573776.013.18.

Amato, Michael S. & Maryellen C. MacDonald. 2010. Sentence processing in an artificial language: Learning and using combinatorial constraints. *Cognition* 116(1). 143–148. https://doi.org/10.1016/j.cognition.2010.04.001.

Arends, Jacques, & Bruyn, Adrienne. 1995. Gradualist and developmental hypotheses. In *Pidgins and creoles: An introduction* (Vol. 15, pp. 111–120). John Benjamins Publishing.

Austin, Keith, Theakston, Anna, Lieven, Elena, & Tomasello, Michael. 2014. Young children's understanding of denial. *Developmental psychology*, *50*(8), 2061.

Baptista, Marlyse. 2002. The syntax of Cape Verdean Creole. The Sotavento varieties. Amsterdam/Philadelphia: John Benjamins.

Baptista, Marlyse, Gelman, Susan A. & Beck, Erica. 2014. Testing the role of convergence in language acquisition, with implications for creole genesis. *International Journal of Bilingualism* 20(3). 269–296.

Baptista, Marlyse. 2016. Creole formation and L2 acquisition: On re-evaluating processes and labels. Guest Column (2). *Journal of Pidgin and Creole Languages*. 31:2, 361-389.

Bardel, Camilla & Falk, Ylva. 2012. The L2 status factor and the declarative/procedural distinction. In Jennifer Cabrelli Amaro, Suzanne Flynn & Jason Rothman (eds.), *Studies in Bilingualism*, vol. 46, 61–78. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/sibil.46.06bar.

Bates, Douglas, Mächler, Martin, Bolker, Ben, & Walker, Steve. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48.

Bickerton, Derek. 1975. *Dynamics of a creole system*. London; New York: Cambridge University Press.

Bickerton, Derek. 1981. *Roots of language*. Ann Arbor: Karoma Publishers.

Bresnan, Joan. 2001. *Lexical-functional syntax*. *Blackwell Textbooks in Linguistics*. Vol. 16.

Brown, Helen, Smith, Kenny, Samara, Anna, & Wonnacott, Elizabeth. 2022. Semantic cues in language learning: an artificial language study with adult and child learners, Language, Cognition and Neuroscience, 37:4, 509-531.

Brown, Richard. 1994. Kresh. *Typological studies in negation*, ed. by Peter Kahrel and René Van Den Berg, 163–189. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.29.09bro.

Burgess, Danielle. 2022. Investigating the presence of NegFirst biases in learning and communication. *Proceedings of the 44th Annual Conference of the Cognitive Science Society*, 1498-1504.

Carpenter, Patricia A., & Just, Marcel A. 1975. Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review, 82*(1), 45–73.

Chang, Suk-Jin. 1996. *Korean*. Amsterdam, Philadelphia: John Benjamins.

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

Clark, Herbert H., & Chase, William G. 1972. On the process of comparing sentences against pictures. *Cognitive Psychology, 3*(3), 472–517

Culbertson, Jennifer. 2012. Typological Universals as Reflections of Biased Learning: Evidence from Artificial Language Learning: Typological Universals as Reflections of Biased Learning. *Language and Linguistics Compass* 6(5). 310–329. https://doi.org/10.1002/lnc3.338.

Culbertson, Jennifer. 2021. Artificial language learning. To appear in Jon Sprouse (ed.), Oxford Handbook of Experimental Syntax. New York, NY: Oxford University Press. lingbuzz/006100

Culbertson, Jennifer, Franck, Julie, Braquet, Guillaume, Navarro, Magda Barrera, & Arnon, Inbal. 2020. A learning bias for word order harmony: evidence from speakers of non-harmonic languages. *Cognition* 204.

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

Dahl, Östen. 1979. Typology of sentence negation. *Linguistics* 17. 79–106. https://doi.org/10.1515/ling.1979.17.1-2.79.

Dahl, Östen. 2010. Typology of negation. In Laurence R. Horn (ed.), *The Expression of Negation*, 9–38. De Gruyter Mouton. https://doi.org/10.1515/9783110219302.9.

Dale, Rick & Lupyan, Gary. 2012. Understanding the origins of morphological diversity: the linguistic niche hypothesis. *Advances in Coplex Systems* 15(3): 1150017-1–1150017-16.

Devos, Maud & van der Auwera, Johann. 2013. Jespersen cycles in Bantu: double and triple negation. *Journal of African Languages and Linguistics*, *34*(2), 205-274.

Dimroth, Christine. 2008. Age effects on the process of L2 acquisition? Evidence from the acquisition of negation and finiteness in L2 German. *Language Learning* 58: 117–150.

Dimroth, Christine. 2010. The Acquisition of Negation. In Laurence R. Horn (ed.), *The Expression of Negation*, 39–72. De Gruyter Mouton. https://doi.org/10.1515/9783110219302.39.

Dryer, Matthew S. 1988. Universals of negative position. In Michael Hammond, Edith A. Moravcsik & Jessica Wirth (eds.), *Typological Studies in Language*, vol. 17, 93. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.17.10dry.

Dryer, Matthew S. 1992. The Greenbergian Word Order Correlations. *Language* 68. 81-138. http://doi.org/10.2307/416370.

Dryer, Matthew S. 2009. The Branching Direction Theory of Word Order Correlations Revisited. *Studies in Natural Language and Linguistic Theory*, 76. 185–207. https://doi.org/10.1007/978-1-4020-8825-4_10.

Dryer, Matthew S. 2013. Negative Morphemes. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) The World Atlas of Language Structures Online. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Fehér, Olga, Ritt, Nikloaus, & Smith, Kenny. 2019. Asymmetric accommodation during interaction leads to the regularisation of linguistic variants. *Journal of Memory and Language* 109, Article 104036.

Fehér, Olga, Wonnacott, Elizabeth & Smith, Kenny. 2016. Structural priming in artificial languages and the regularisation of unpredictable variation. *Journal of Memory and Language* 91. 158–180. https://doi.org/10.1016/j.jml.2016.06.002.

Feiman, Roman, Mody, Shilpa, Sanborn, Sophia & Carey, Susan. 2017. What Do You Mean, No? Toddlers' Comprehension of Logical "No" and "Not." *Language Learning and Development* 13(4). 430–450. https://doi.org/10.1080/15475441.2017.1317253.

Ferreira, Victor S. 2019. A mechanistic framework for explaining audience design in language production. *Annual Review of Psychology* 70, 29-51.

Frege, Gottlob. 1892. On Sinn and Bedeutung. In M. Beaney (ed.), *The Frege Reader,* 151–71. Oxford: Blackwell.

Giuliano, Patricia. 2003. Negation and relational predicates in French and English as second languages. In Christine Dimroth and Marianne Starren (eds.), *Information Structure and the Dynamics of Language Acquisition*, 119–158. Amsterdam: Benjamins.

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

Gruzdeva, Ekaterina. 2001. Imperative sentences in Nivkh. In Viktor S. Xrakovskij (ed.), *Typology of imperative constructions*, 59–77. Lincom Europa.

Güldemann, Tom. 2011. Proto-Bantu and Proto-Niger-Congo: Macro-areal typology and linguistic reconstruction. In Osamu Hieda, Christa König, and Hirosi Nakagawa (eds.), *Geographical typology and linguistic areas*, 109–42. Amsterdam: John Benjamins.

Harris, Martin. 1978. *The evolution of French syntax: a comparative approach* (Longman Linguistics Library 22). London; New York: Longman.

Haspelmath, Martin and the APiCS Consortium. 2013. Position of standard negation. In Michaelis, Susanne Maria, Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds.). *The atlas of pidgin and creole language structures.* Oxford: Oxford University Press.

Hawkins, John A. 1994. *A performance theory of order and constituency.* Cambridge University Press.

Hawkins, John A. 2004. *Efficiency and complexity in grammars.* Oxford: Oxford University Press.

Heine, Bernd, & Kuteva, Tania. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.

Holm, John. 1988. *Pidgins and creoles. Vol 1. Theory and structure*. Cambridge: Cambridge University Press.

Horn, Laurence R. 1989. *A natural history of negation*. Chicago: Univ. of Chicago Pr.

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). 815–821. https://doi.org/10.1037/a0015097.

Hudson Kam, Carla L., & Newport, Elissa L. 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., & Newport, Elissa L. 2009. Getting it right by getting it wrong: When learners change languages. *Cognitive psychology*, 59(1), 30–66.

Hyltenstam, Kenneth. 1977. Implicational patterns in interlanguage syntax variation. *Language Learning* 27(2). 383–410. https://doi.org/10.1111/j.1467-1770.1977.tb00129.x.

Jeon, K. Seon. 2014. Language Processing Difficulty in the Development of Negation : the Case of L2 Korean Negation. *Language Information*, 19, 145-166.

Jespersen, Otto. 1917. *Negation in English and other languages*. Copenhagen: A. F. Høst.

Kanagy, Ruth. 1994. Developmental Sequences in Learning Japanese: A Look at Negation. *Issues in Applied Linguistics* 5. https://doi.org/10.5070/l452005197.

Kaup, Barbara, Zwaan, Rolf A., & Lüdtke, Jana. (2006). Procesing negated sentences with contradictory predicates: Is a door that is not open mentally closed? *Journal of Pragmatics, 38*, 1033–1050.

Kaup, Barbara, Yaxley, Richard H., Madden, Carol J., Zwaan, Rolf A., & Lüdtke, Jana. 2007. Experiential simulations of negated text information. *Quarterly Journal of Experimental Psychology, 60*, 976–990.

Keogh, Aislinn, Simon Kirby & Jennifer Culbertson. 2022. In search of a unified mechanism for regularisation of linguistic variation. *Proceedings of the 23rd Amsterdam Colloquium*, 360-366.

Kerswill, Paul & Williams, Ann. 2000. Creating a new town koine: Children and language change in Milton Keynes. *Language in Society*, 29. 65-115.

Kim, Boomee. 1991. The acquisition of the two negation forms in Korean. *Harvard Studies in Korean Linguistics* 4. 13–25.

Kim, Young-Joo. 1997. The acquisition of Korean. In Slobin, D. (ed.), *The crosslinguistic study of language acquisition*, Vol 4, 335-443. Hillsdale, NJ: Lawrence Erlbaum.

Kim, Hoe Kyeung & Hongoak Yun. 2013. Examining the Influence of L1 on Learning Korean Negation Forms. *Foreign Languages Education* 20. 415–434.

Kiparsky, Anthony. 1996. The shift to head-initial VP in Germanic. In H. Thrainsson, J. Peter, and S. Epsitein (eds.), *Comparative Germanic Syntax*. Kluwer.

Kroch, Anthony, Taylor, Ann & Ringe, Donald. 2000. The Middle English verb-second constraint: A case study in language contact and language change. In S. C. Herring, P. van Reenen, and L. Schøsler (eds.), *Textual parameters in older languages*, pp. 353–391. Amsterdam: John Benjamins.

Kuznetsova, Alexandra, Brockhoff, Per B., Christensen, Rune H. B. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82(13), 1-26.

Langdon, Margaret H. 1970. *A grammar of Diegueño: The Mesa Grande Dialect*. Berkeley: University of California Press.

Larsen-Freeman, Diane, & Long, Michael. 1994. *An Introduction to Second Language Acquisition Research*. London: Routledge.

Lee, Jin-Hwa. 2008. A Corpus-based Study of Short-form and Long-form Negations in Korean. *The Korean Language in America* 13. 43–57.

Lee, Miseon. 2017. Incremental processing of negation: Evidence from Korean. *Linguistic Research* 34. 205–224.

Lee, Sunyoung. 2012. Distribution of Sentential Negation in Korean: A Corpus-based View. *Studies in Modern Grammar* 67.107–123.

Lehmann, Winfred P. 1973. A structural principle of language and its implications. *Language*, 49, 47-66.

Lenth, Russell V. 2020. *emmeans: Estimated Marginal Means, aka Least-Squares Means*. R package version 1.5.0.

Levelt, Willem J. M. 1989. *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Lightfoot, David. 1997. Catastrophic change and learning theory. *Lingua* 100, 171–192.

Lipski, John. 2018. Palenquero vs. Spanish negation: Separate but equal? *Lingua* 202, 44–57. https://doi.org/10.1016/j.lingua.2017.12.007.

Lupyan, Gary, & Dale, Rick. 2010. Language structure is partly determined by social structure. PloS one, 5, e8559.

MacDonald, Maryellen. C. 2013. How language production shapes language form and comprehension. *Frontiers in Psychology* 4, Article 226.

MacDonald, Maryellen. C., & Just, Marcel A. 1989. Changes in activation levels with negation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(4), 633–642.

Maldonado, Mora & Culbertson Jennifer. 2021. Nobody Doesn't Like Negative Concord. *Journal of Psycholinguist Research*;50(6):1401-1416. doi:10.1007/s10936-021-09816-w.

Martin, Alexander & Culbertson, Jennifer. 2020. Revisiting the suffixing preference: Native language affixation patterns influence perception of sequences. *Psychological Science* 31(9). 1107–1116.

Maurer, Philippe and the APiCS Consortium. 2013. Position of tense, aspect, and mood markers in relation to the verb. n Michaelis, Susanne Maria, Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds.), *The atlas of pidgin and creole language structures.* Oxford: Oxford University Press.

Meisel, Jürgen. 1997. The acquisition of the syntax of negation in French and German: contrasting first and second language development. *Second Language Research* 13: 227–263.

Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds.) 2013. *Atlas of Pidgin and Creole Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Miestamo, Matti. 2005. *Standard Negation*. Berlin, New York: Walter de Gruyter. doi:10.1515/9783110197631.

Mous, Maarten. 1992. *A grammar of Iraqw.* Doctoral dissertation, Leiden University.

Mufwene, Salikoko. 1996. The Founder Principle in creole genesis. Diachronica 13: 83–134.

Mufwene, Salikoko. 2001. The Ecology of Language Evolution. Cambridge: Cambridge University Press.

Mufwene, Salikoko S. 2008. Language evolution: contact competition, and change. London: Continuum Press.

Nam, Yunju. 2016. An ERP study on the processing of Syntactic and lexical negation in Korean. *Korean Journal of Cognitive Science* 27(3), 469-499.

Nieuwland, Mante S., & Kuperberg, Gina R. 2008. When the Truth Is Not Too Hard to Handle: An Event-Related Potential Study on the Pragmatics of Negation. *Psychological Science*, *19*(12), 1213–1218.

Pak, Marjorie. 2020. Clause-final negation and the Jespersen cycle in Logoori. *Proceedings of the Linguistic Society of America*, 187–199.

Park, Heeheon. 1998. Acquisition of Negation in Korean. *Korean Linguistics* 9.111–131. doi:10.1075/kl.9.04hp.

Payne, John R. 1985. Negation. In Timothy Shopen (ed.), *Language typology and syntactic description: Vol. 1. Clause structure*, 197–242. Cambridge: Cambridge University Press

Paloposki, Päivi. 1999. *Puolan kielioppi [Polish grammar]*. Helsinki: Finn Lectura.

Pickering, Martin J. & Garrod, Simon. 2004. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences* 27(02). https://doi.org/10.1017/S0140525X04000056.

Pienemann, Manfred. 1998. *Language processing and second language development: Processability Theory*. Amsterdam: John Benjamins.

Pienemann, Manfred. 2000. Psycholinguistic Mechanisms in the Development of English as a Second Language. In *Language Use, Language Acquisition and Language History:*

(Mostly) Empirical Studies in Honour of Rüdiger Zimmermann, edited by Ingo Plag and Klaus Schneider, 99–118. Trier: Wissenschaftlicher Verlag Trier.

Pienemann, Manfred. 2005. An introduction to Processability Theory. *Cross-Linguistic Aspects of Processability Theory*, ed. by Manfred Pienemann, 1–60. John Benjamins Publishing Company. https://doi.org/10.1075/sibil.30.03pie.

Pienemann, Manfred. 2015. An Outline of Processability Theory and Its Relationship to Other Approaches to SLA. *Language Learning* 65.123–151. https://doi.org10.1111/lang.12095.

Plag, Ingo. 2008. Creoles as interlanguages: Syntactic structures. *Journal of Pidgin and Creole Languages* 23(2). 307–328. https://doi.org/10.1075/jpcl.23.2.06pla.

R Core Team. 2021. R: A language and environment for statistical computing. *R Foundation for Statistical Computing*. Vienna, Austria.

Rhee, Seonghwa. 2003. When Stop Goes Further: From Malta Stop to Auxiliary Verbs in Korean. *Korean Language Research* 13: 309-339.

Roberts, Gareth & Sneller, Betsy. 2020. Empirical foundations for an integrated study of language evolution. *Language Dynamics and Change* 10(2), 188–229.

Saldana, Carmen, Oseki, Yohei, & Culbertson, Jennifer. 2021. Cross-linguistic patterns of morpheme order reflect cognitive biases: An experimental study of case and number morphology. *Journal of Memory and Language* 118.

Saldana, Carmen, Smith, Kenny, Kirby, Simon. & Culbertson, Jennifer. 2021. Is Regularization Uniform across Linguistic Levels? Comparing Learning and Production of Unconditioned Probabilistic Variation in Morphology and Word Order. *Language Learning and Development*, 17(2), 158-188.

Samara, Anna, Smith, Kenny, Brown, Helen, & Wonnacott, Elizabeth. 2017. Acquiring variation in an artificial language: Children and adults are sensitive to socially conditioned linguistic variation. *Cognitive Psychology* 94. 85–114. https://doi.org/10.1016/j.cogpsych.2017.02.004.

Schneider, Edgar W. 2000. Feature diffusion vs. contact effects in the evolution of new Englishes: A typological case study of negation patterns. *English World-Wide* 21, 201–230.

Schwab, Jessica F., Lew-Williams, Casey, & Goldberg, Adele E. When regularization gets it wrong: children over-simplify language input only in production. J Child Lang. 2018 Sep;45(5):1054-1072. https://doi.org/10.1017/S0305000918000041.

Schwegler, Armin. 1983. Predicate negation and word-order change: A problem of multiple causation. *Lingua* 61, 297–334.

Silva-Corvalán, Carmen. 2008. "The Limits of Convergence in Language Contact." *Journal of Language Contact* 2. 213–24.

Snyman, Jan Winston. 1970. *An introduction to the !Xũ (!Kung) language*. Cape Town: Balkema.

Stanhope, John M. 1980. *The Language of the Rao People, Grengabu, Madang Province, Papua New Guinea*. Canberra: Australian National University.

Strawson, Peter F. 1950. On referring. *Mind*, 59: 320–44.

Tian, Ye & Breheny, Richard. 2019. Pragmatics and negative sentence processing. In Chris Cummins & Napoleon Katsos (eds.), *Handbook of experimental pragmatics*. Oxford, UK: Oxford University Press.

Tian, Ye, Breheny, Richard, & Ferguson, Heather. 2010. Why we simulate negated information: A dynamic pragmatic account. *Quarterly Journal of Experimental Psychology,* 63(12), 2305–2312.

Tily, Harry & Jaejer, Florian T. 2011. Complementing quantitative typology with behavioral approaches: Evidence for typological universals. *Linguistic Typology* 15(2). https://doi.org/10.1515/lity.2011.033.

Todd, Loreto. 1991. *Talk Pidgin: A Structured Course in West African Pidgin English*. Leeds: Tortoise Books.

Toribio, Almeida Jacqueline. 2004. Convergence as an optimization strategy in bilingual speech: Evidence from code-switching. *Bilingualism: Language and Cognition* 7(2). 165–173. https://doi.org/10.1017/S1366728904001476.

Tottie, Gunnel. 1991. *Negation in English speech and writing: a study in variation* (Quantitative Analyses of Linguistic Structure Series). San Diego: Academic Press.

van der Auwera, Johan. 2006. Why languages prefer prohibitives. *Journal of Foreign Languages* 1: 2–25.

van der Auwera, Johan and Ludo Lejeune (with Valentin Goussev). 2013. The Prohibitive. In Dryer, Matthew S. & Haspelmath, Martin (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

van Gelderen, Elly. 2008. Negative cycles. *Linguistic Typology* 12(2), 195–243.

Van Olmen, Daniël. 2010. Typology meets usage: The case of the prohibitive infinitive in Dutch. *Folia Linguistica* 44(2), 471–508.

Van Olman, Daniël. 2021. On order and prohibition. *Studies in Language* 45(3), 520–556.

Vennemann, Theo. 1974. Topics, subjects and word order. From SXV to SVX via TVX. In *Historical linguistics*, vol. 1, 339–376. Amsterdam: North-Holland.

Vidal, Alejandra. 2001. *Pilagá grammar (Guaykuruan family, Argentina)*. Doctoral dissertation, University of Oregon.

Voeten, Cesko, C. 2022. *buildmer: Stepwise Elimination and Term Reordering for Mixed-Effects Regression.* R package version 2.7, <https://CRAN.R-project.org/package=buildmer>.

Wason, Peter Cathcart. 1965. The contexts of plausible denial. *Journal of verbal learning and verbal behavior* 4: 7–11.

Wason, Peter Cathcart. 1972. In real life negatives are false. *Logique et Analyse* 15: 17-38.

Weir, E.M. Helen. 1994. Nadëb. *Typological studies in negation*, ed. by Peter Kahrel and René van den berg, 291–323. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.29.14wei.

Willis, David. 2010. Motivating the emergence of new markers of sentential negation: the case of Welsh *ddim*. *Diachronica* 27, 110–156.

Willis, David. 2016. Incipient Jespersen's cycle in Old English negation. In Sten Vikner, Henrik Jørgensen & Elly van Gelderen (eds.), *Let us have articles betwixt us – Papers in Historical and Comparative Linguistics in Honour of Johanna L. Wood,* 465-491.

Wode, Henning. 1981. *Learning a second language.* Tübingen: Narr.

Zeijlstra, Hedde. 206. Diachronic developments in the domain of negation. *Language and Linguistics Compas* 10(6), 284–295.