2025. Proc Ling Soc Amer 10(1). 5955. https://doi.org/10.3765/plsa.v10i1.5955

Pattern deduction in linguistically attested and unattested grammars

Sadhwi Srinivas, Kaitlyn Harrigan, Aidan Burnham, Nicholas Voivoda & Berit Wilkins*

Abstract. A popular hypothesis in linguistics posits that language learners are biologically predisposed to learn structures attested in human language – for example, a hierarchically nested phrase structure, while eschewing hypotheses for linguistically unattested structures – for example, one consisting of non-consecutive, linearly alternating "constituents". The current study explores the robustness of such a predisposition within a controlled artificial language learning task as well as a non-linguistic, general puzzle-solving task. We find evidence that suggests learners more easily acquire linguistically attested hierarchically structured patterns compared to unattested non-hierarchical ones within the non-linguistic task, but not within the language task. We discuss the puzzling nature of this finding, and some work in progress to further unravel the source of this result.

Keywords. hierarchical bias in language learning; pattern deduction; hierarchically structured grammars; artificial language learning; learnability

1. Introduction. Language is acquired through a mix of language-specific and domain-general cognitive processes. Work in language acquisition has aimed to identify how the learner utilizes each of these types of processes. In particular, as a consequence of the famous "Poverty of Stimulus" argument (Chomsky 1986), language learners have long been hypothesized to have a bias towards acquiring hierarchically organized phrase structure rules over arguably simpler, albeit linguistically unattested, linear generalizations. For example, English-learning children's acquisition of the auxiliary fronting rule in yes/no questions is often cited as evidence for such a hierarchical bias. Consider examples (1)-(2) below. The rule for obtaining the yes/no question in (1b) from the assertion in (1a) can be stated as a hierarchical generalization ("move the main clause auxiliary to the front") or a linear one ("move the linearly precedent auxiliary to the front"). However, the more complex example (2) clarifies that only the hierarchical structure-based generalization is successful (2b); attempting to front the linearly precedent auxiliary leads to ungrammaticality (2c). Children are claimed to seldom receive disambiguating datapoints such as (2) in their input, yet they never produce ungrammatical structures of the kind in (2c). This has lead to the influential hypothesis that infants have an innate bias for hierarchical generalizations – that is, they assume hierarchical structure when deducing language rules.

(1) a. The man <u>is</u> a fool.
    b. <u>Is</u> the man a fool?

(2) a. The man who is a fool <u>is</u> amusing.

* We thank the Linguistics program, the English department, and the Charles Center at William & Mary for contributing funding for this work. We are also grateful to the participants of the linguistics research group meetings, and the reviewers and audiences at LSA 2025 for their support and valuable feedback.
Authors: Sadhwi Srinivas, W&M (ssrinivas@wm.edu), Kaitlyn Harrigan, W&M (kharrigan@wm.edu), Aidan Burnham, W&M (aeburnham@wm.edu), Nicholas Voivoda, W&M (nwvoivoda@wm.edu) & Berit Wilkins, W&M (bfwilkins@wm.edu).

b. <u>Is</u> the man who is a fool amusing?
c. *Is the man who a fool <u>is</u> amusing?

In the current study, we seek to further explore the source of this purported underlying bias towards hierarchical structure during language acquisition – particularly, the question of whether such a bias is language-specific, or whether it is a more general cognitive bias. Let us call the former the *domain-specific hypothesis* for language learning, first suggested by Chomsky and standardly assumed in mainstream generative linguistics (e.g., Yang 2006). Let us call the latter the *domain-general hypothesis* for language learning which finds support under the general umbrella of usage-based linguistics (e.g., Goldberg 2003, Tomasello 2003, Bybee 2010), positing that language learning proceeds via general cognitive mechanisms subject to general cognitive pressures that also aid learning in other kinds of contexts. Under this type of view, the observation that only certain types of patterns are attested in human languages and entertained by the learner would be explained by appealing to more widely applicable constraints on human learning. In other words, only those patterns are expected to show up in language that are generally able to be learned, and linguistically-unattested patterns would be unattested in other domains as well. To adjudicate between these two hypotheses, we examine the learnability of two types of grammars in the current study (one hierarchically structured and one non-hierarchically structured grammar) in two types of tasks (a language learning task and a non-linguistic, general puzzle-solving task).

The first part of our study seeks to investigate the learner's biases within a language learning context. To do so, we employ an artificial language learning design that builds closely on Takahashi & Lidz (2007), in which adult participants were exposed to constructed, hierarchically structured grammars over a series of learning sessions. At the end of the learning procedure, participants in Takahashi & Lidz's study were found not only to have learned the hierarchical constituency structure, but also to have generalized beyond what was observed during the learning stage to conclude that the hierarchically organized constituents could undergo various kinds of transformations. In one of our study conditions, we employ a similar hierarchically structured, constructed grammar as a way to replicate a part of what these authors found – that hierarchical structures are indeed learnable within a language learning context.

In addition to a hierarchical grammar, we also employ an additional non-hierarchical grammar in this task, in order to verify that the former is in fact learned more easily than the latter. Empirical support for the unlearnability of non-hierarchical patterns within a language learning context is much sparser in the literature. This is most directly tested in Smith et al. (1993), where a conlang *Epun* was used to examine the ease of acquisition of various rules by a language savant named Christopher as well as four undergraduate controls. One of the rules in *Epun* was a non-hierarchical "structure-independent" rule, in which an emphatic morpheme *nog* was governed by an "arithmetically-determined" computation rather than a hierarchical structurally-determined computation. Specifically, this emphatic element always attached to the linearly third orthographic word in a given sentence, regardless of the lexical category of the word or the constituency structure within the sentence – as shown in (3).[1]

(3) a. *Fa* *zaddil-in* *ha-bol-u-nog* *guv.*
The man-NOM PAST-go-3MS-EMPH yesterday
"The man did go yesterday." (Smith et al. 1993, ex. 42a)

b. *Lodon-in* *ha-bol-u* *guv-nog.*
Lodon-NOM PAST-go-3MS yesterday-EMPH
"Lodon did go yesterday." (Smith et al. 1993, ex. 42b)

Smith et al. report that the rule for *nog* was essentially unlearnable both by Christopher as well as the four undergraduate controls – though they do not include any quantitative results. While this result is potentially highly theoretically relevant, we believe its strength is limited by a couple of considerations. First, the low number of control participants makes it difficult to gauge if this result truly generalizes across all contexts of language learning. Second, and more importantly, the structure-independent rule in this study was imposed on a grammar that was underlyingly hierarchically-structured; most other rules in *Epun* were defined in a structure-dependent way. This opens up the possibility that participants were only resistant to generalizing an arithmetically-computed rule on top of an underlying hierarchical grammar, and crucially, that they would be more successful at generalizing such a rule within a system where all rules are similarly arithmetically-computed. A strong version of the hypothesis that human learners only entertain hierarchical structure-dependent rules within a language learning context would predict that structure-independent rules are more widely unlearnable, and not just as part of a system that otherwise contains hierarchical structure-dependent rules. It is this stronger prediction that we wish to test in the current study, with a larger set of participants in a more controlled fashion.

In the second part of our study, we wish to compare results from the language task to pattern learning and generalization outside of a language learning situation, in a general puzzle-solving context, as this can shed light on what kinds of patterns are learned due to language-specific biases *vs.* general cognitive biases. If the preference for learning hierarchical structure-dependent patterns over structure-independent ones is observed only in the language learning task but not in the general puzzle-solving task, this would suggest that it is a language-specific bias. More broadly, it provides evidence for language-specific mechanisms that are separate from non-language ones, lending support to the domain-specific hypothesis. On the other hand, if the hierarchical preference spans across the language learning and the general puzzle-solving tasks, that would suggest that the preference for generalizing to hierarchical patterns is domain-general.

In Section 2, we turn to describing the overall set-up of our study and the findings in more detail. Section 3 then discusses the overall implications of our findings, and steps for future work. Section 4 concludes.

[1] If a sentence contained fewer than three orthographic words, the emphatic element *nogin* was added to the final word in the sentence. Smith et al. report that Christopher had special difficulty with *nogin*.

2. Experiment: Pattern deduction in language and non-language contexts. Participants completed a pattern deduction task, in which they attempted to "escape from an escape room." To successfully complete the task, they needed to learn a set of patterns that they were exposed to in a series of phases showing examples of acceptable strings.

2.1. PARTICIPANTS. Participants were 125 undergraduates recruited from William & Mary's introductory Linguistics and Psychology courses. All participants received course credit for their participation in the study.

2.2. DESIGN. The current study utilizes a 2x2 design. Participants were randomly assigned to an escape room *context* (*language* or *shapes*). In both contexts, participants were told that their task was to escape from a virtual escape room. In the *language context*, participants were told that they needed to learn some basic words and structures of a language, "Nog", in order to communicate with the guards to make their escape. In the *shapes context*, participants were told that their task was to learn what sequences of shapes would successfully unlock a series of lock-boxes to facilitate their escape. Participants were shown sequences of either words (*language context*) or shapes (*shapes context*) that were consistent with one of two possible *patterns*: one that is linguistically attested (*hierarchical pattern*), and one that is systematic but linguistically **un**attested (*linear pattern*). Throughout the study, participants were asked to use their knowledge of the patterns based on the examples they saw during the exposure phases to choose between novel sequences in a forced choice task.

Basic vocabulary. Learning began in each context with a basic "vocabulary". In the *language* context, participants were told that these were basic sentences of the language; in the *shapes* context, participants were told that they were learning the basic sequences of shapes to unlock the first level of lockboxes. The acceptable sequences for each context's vocabulary was the same: they consisted of five positions, each of which had two possible words or shapes, leading to 32 possible grammatical basic sequences. Table 1 shows all words or shapes used for each of the possible positions. The shapes we used were adapted from Fiser & Aslin (2001).

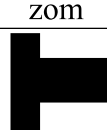
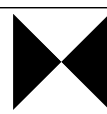
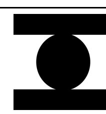
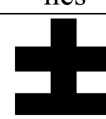
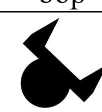
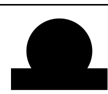
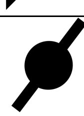
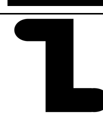
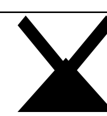
		A	B	C	D	E
language	option 1	nok	zin	sel	kan	tal
	option 2	zom	ket	biv	hes	bop
shapes	option 1					
	option 2					

Table 1. Vocabulary items for *language* and *shapes* contexts.

Hierarchical pattern. The *hierarchical* pattern was a hierarchically organized "grammar" consisting of three phrases, each containing two or three elements (Figure 1). This grammar included hierarchical embedding, with one phrase (CP) embedded inside another phrase (EP). It was intended to be a simplified version of natural language, in which constituents are consecutive strings of words which may be hierarchically embedded, and over which transformations occur.

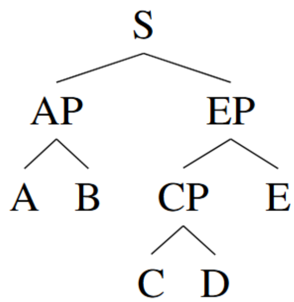


Figure 1. Constituents of the *hierarchical* grammar.

Two operations occurred in the transformed sentences of this grammar: deletion and movement. In the deletion sentences, one of the three constituents was deleted. In the movement sentences, one of the three constituents was moved to the beginning of the sequence. We did not include transformed movement sequences for the constituent [AB], as this was already at the beginning of the utterance in the base form. Table 2 shows all possible transformed sequences.

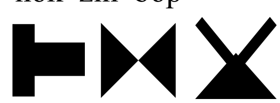
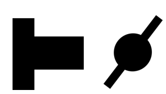
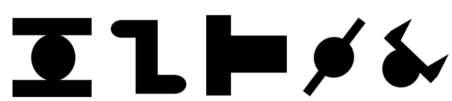
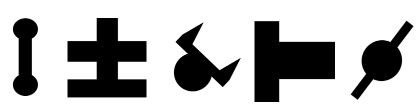
	constituent	sequence	examples
deletion	AB	CDE	sel hes bop 
	CD	ABE	nok zin bop 
	CDE	AB	nok ket 
movement	~~AB~~	~~ABCDE~~	
	CD	[CD]ABE	sel hes nok ket tal 
	CDE	[CDE]AB	biv kan tal nok ket 

Table 2. Transformed sequences in the *hierarchical* grammar.

Linear pattern. The linguistically unattested *linear* pattern was a systematic pattern that also utilized transformations over "constituents," but in this case the constituents were not consecutive strings, but alternating words/shapes. The grammar consisted of two constituents: words/shapes A, C & E, and words/shapes B & D (Figure 2).

A B C D E

Figure 2. Constituents of *linear* grammar.

The same two operations occurred in the transformed sentences of this grammar: deletion and movement. In the deletion sentences, one of the two constituents was deleted. In the movement sentences, one of the two constituents would move to the beginning of the sequence. Table 3 shows all possible transformed sequences.

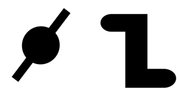
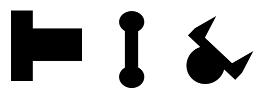
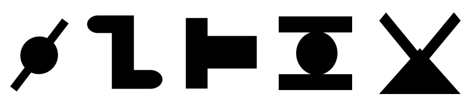
	constituent	sequence	examples
deletion	ACE	BD	ket hes 
	BD	ACE	nok biv tal 
movement	ACE	[ACE]BD	zom biv tal zin hes 
	BD	[BD]ACE	ket hes nok sel bop 

Table 3. Transformed sequences in the *linear* grammar.

2.3. PROCEDURE. The study was deployed via Qualtrics and guided participants through the multiple exposure and test phases. Participants were randomly assigned one of the four possible conditions: *hierarchical language*, *linear language, hierarchical shapes,* or *linear shapes.*

Vocabulary. In an initial vocabulary learning phase, participants were exposed to all 32 possible basic combinations. They were then given eight vocabulary check questions, which asked them to choose which of two combinations was one they had been exposed to, testing their ability to distinguish grammatical from ungrammatical basic strings. The ungrammatical foils were scrambled sequences, where the words or shapes were not in the correct order. This vocabulary check was then repeated, for a total of 64 examples sequences (each possible combination twice) and 16 vocabulary check questions.

Transformations. Participants then moved on to the transformation exposure phases, where they were told that they were going to see some harder sequences. During this phase, they were exposed to a set of 74 transformed sequences, including a representative subset of possible sequences with constituents deleted (21 sequences) or moved (53 sequences). The examples included sequences of both transformation types for each constituent. Participants were then given eight test questions, in which they were asked to choose which of two sequences was more likely to be a grammatical sentence of the target language, or a viable combination of shapes. They were then exposed to the same 74 transformed sequences again and given another eight test questions. Each participant responded to a total of 16 test questions: 5 deletion and 11 movement sequences. The ungrammatical sequences in both conditions consisted of an ungrammatical sentence in which a transformation of a basic sentence was generated by deleting or moving a non-constituent but consecutive string. The transitional probabilities between consecutive elements within the transformed grammatical sequences across both conditions were controlled to be roughly equivalent, as were the transitional probabilities for the ungrammatical foils. This was done to ensure that any observed learning biases could be attributed to the attested *vs.* unattested natures of the grammars alone and not to any type of lower-level statistical cues.

2.2. RESULTS. We excluded participants who chose the grammatical string at a rate less than 75% in the vocabulary check phase (n=5), leaving us with 120 participants for the final analysis—30 in each of four conditions. All groups had average vocabulary scores above 95% (Table 4).

language		*shapes*	
hierarchical (n=30)	*linear* (n=30)	*hierarchical* (n=30)	*linear* (n=30)
95.0%	95.6%	95.8%	95.4%

Table 4. Average vocabulary scores by condition.

Overall, we find that participants learned at above chance levels in all four conditions. They successfully learned both the *hierarchical* and *linear* patterns in both *language* and *shapes* contexts (Table 5, Figure 3). In both contexts, participants learned the *hierarchical* pattern better than the *linear* pattern. Additionally, they showed higher accuracy overall in the *shapes* context compared to the *language* context.

language		*shapes*	
hierarchical (n=30)	*linear* (n=30)	*hierarchical* (n=30)	*linear* (n=30)
69.6%	63.5%	81.0%	66.0%

Table 5. Average test scores by condition.

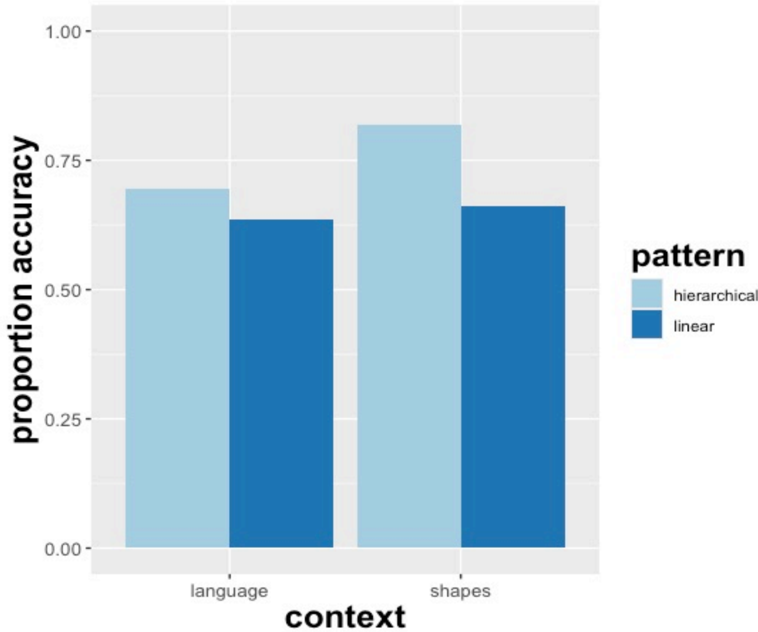


Figure 3. Accuracy on test items by condition.

The results were analyzed in R (R Core Team 2021, RStudio Team 2020) using a binary logistic regression model with *accuracy* as the dependent variable, *context* (*language* v. *shapes*) and *pattern* (*hierarchical* v. *linear*) as fixed effects, and *subject* as a blocking variable. We find a main effect of *context* [$X^2_{(1)} = 6.16$, $p = 0.01$], but no main effect of *pattern* [$X^2_{(1)} = 1.59$, $p = 0.21$] and no interaction [$X^2_{(1)} = 1.87$, $p = 0.17$]. To compare performance across the two *patterns* within each *context*, we looked at pairwise comparisons. We find a significant difference between *hierarchical* and *linear* patterns in the *shapes* context [$p < 0.05$] but not in the *language* context [$p = 0.21$]. Participants were significantly better at learning the *hierarchical* pattern compared to the *linear* pattern in the *shapes* context, but not in the *language* context.

2.3 EXPERIMENT DISCUSSION. Our results indicate that while participants learned constituents successfully (above chance) in the linguistically attested *hierarchical* grammar as well as the linguistically unattested but "logical" *linear* grammar in both linguistic and non-linguistic learning contexts, they do not learn equally well in all conditions. We find a numerical but non-significant advantage to the hierarchical grammar overall, but a significant advantage for learning the hierarchical pattern in the *shapes* condition. This finding is surprising for two reasons. First, the findings are consistent with a hierarchical bias that is actually stronger in non-linguistic contexts than in linguistic contexts. This is unexpected given the vast literature pointing to a domain-specific hierarchical bias in language contexts. Second, claims have also been made that linguistically unattested patterns are not considered in language learning contexts. We find that participants are also learning the linguistically unattested pattern above chance levels. At this point it remains unclear whether these findings are due to task effects, or reflective of biases in learning patterns in and outside of linguistic contexts.

3. General Discussion. In this study, we sought to further investigate a widely reported result in the linguistics and language learning literature, that language learners bring in an inductive bias to the process of learning that allows them to pick up linguistically attested hierarchical patterns more easily than linguistically unattested ones. Our study had two broad goals. The first goal was to replicate and strengthen the results from previous work that support the existence of a hierarchical bias in language learning. Specifically, we hypothesized that adult language learners would be good at learning hierarchical structure-dependent rules (replicating Takahashi & Lidz 2007) but significantly worse at learning a linguistically unattested, structure-independent pattern (replicating and strengthening Smith et al. 1993). A second main goal of the study was to further explore the source of the preference for hierarchical structure-dependent rules in language learning: whether it is a domain-specific bias, in which case a preference for hierarchical patterns would not be expected to extend to non-linguistic tasks, or if it is a domain-general bias, in which case we might expect the hierarchical preference to extend to other tasks as well. To address these goals, we tested participants' ability to learn the same linguistically attested and unattested rules within a language learning context and a non-language learning, general puzzle-solving context.

Overall, participants showed the ability to successfully learn both linguistically attested and unattested patterns, within both the language learning context and the puzzle-solving context. While it is expected – given previous work (Takahashi & Lidz 2007) – that participants would be able to learn the linguistically attested patterns, the fact that they were also successful with the unattested pattern contradicts theoretical expectations (cf. Smith et al. 1993). Moreover, we observed no significant differences in how well they learned the language-unattested pattern compared to the language-attested one – a result inconsistent with the influential claim that language learners discard non-hierarchically structured patterns from their hypothesis space right from the outset. What is further surprising is that we did find evidence for a structure-dependent bias in the general puzzle-solving task, in which participants were significantly more successful at learning the hierarchical, linguistically attested rule than the linguistically unattested one.

This pattern of findings does not receive a straightforward theoretical explanation. If the preference for the structure-dependent pattern in the non-linguistic task is due to a domain-general bias, this does not explain why a similar preference is not observed within the language task. But given the result in the non-linguistic task, we also cannot conclude the total lack of a hierarchical bias in cognition. There is moreover no precedence to expect such a preference only outside of a language learning/processing context. In light of this, we believe the simplest, most likely explanation for why a preference for structure-dependent operations was not observed in the language task boils down to issues with the way that our language task was set up. Below, we discuss a few of these issues along with some work that is underway to resolve them.

The first, main issue may have been simply with the set-up of our language task. Recall that participants are asked to "communicate" with the Nog-speaking guards guarding the escape room. However, because the training phase only instructs participants in possible structures in Nog but does not instruct them in any form-to-meaning mappings, participants in fact have no access to any meanings. This means that a true communication task, wherein any meaningful information is exchanged between the interlocutors, is impossible. This issue may be resolved in one of two ways – either by adding form-to-meaning mappings, or by modifying the instructions so that the task is no longer a communicative one where participants are actually expected to convey any meaningful information. Implementing the former introduces a fundamental change to the experimental paradigm, whereas our current goal has been merely to test structural generalizations. As such, in a follow-up study, we are testing a version of the task with modified instructions that removes any communicative expectations. In this version, participants are merely asked to pretend to be speakers of Nog, so that the Nog-speaking guards can be tricked into thinking that they are indeed one of them and letting them pass through the doors.

Aside from the above, we additionally note that the vocabulary items in the language task were very phonotactically simple and English-like, which could have led to participants reading them aloud and clicking through the survey without necessarily having internalized them. No such shortcuts were possible with the shapes. Finally, the visual set-up within the language task may have been overall less compelling than the non-language task, which may have led to decreased overall task attention. The follow-up version of the study will resolve both of these issues as well – by introducing non-English-like vocabulary items that are harder to sound out and by upgrading the quality of the visual set-up, respectively.

4. Conclusion. This study attempts to systematically test a well-known hypothesis in the literature – that language learners are predisposed to learn structure-dependent patterns that are attested in human languages, eschewing hypotheses for unattested structures, and moreover that such a predisposition is language-specific. The results were surprising. While we did not observe a strong preference for learning structure-dependent patterns in the language task, we did find such a preference in the non-language task. Follow-ups are underway to get to the bottom of this result.

References

Bybee, Joan. 2010. *Language, usage and cognition.* Cambridge: Cambridge University Press.

Chomsky, Noam. 1986. *Knowledge of language*. New York: Praeger.

Fiser, József & Richard. N. Aslin. 2001. Unsupervised Statistical Learning of Higher-Order Spatial Structures from Visual Scenes. *Psychological Science*, 12(6). https://doi.org/10.1111/1467-9280.00392

Goldberg Adele. 2003. Constructions: a new theoretical approach to language. *Trends in cognitive sciences* 7(5). https://doi.org/10.1016/s1364-6613(03)00080-9

Takahashi, Eri & Jeffrey Lidz. 2007. Beyond statistical learning in syntax. In *Language acquisition and development: Proceedings of GALA* 446.

R Core Team. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.r-project.org/

RStudio Team. 2020. RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

Smith, Neil, Ianthi-Maria Tsimpli, & Jamal Ouhalla. 1993. Learning the impossible: The acquisition of possible and impossible languages by a polyglot savant. *Lingua* 91(4). https://doi.org/10.1016/0024-3841(93)90002-E.

Tomasello, Michael. 2003. *Constructing a language: A usage-based theory of language acquisition.* Cambridge, MA: Harvard University Press.

Yang, Charles. 2006. *The Infinite Gift: How Children Learn and Unlearn the Languages of the World.* New York: Scribner.