



Commentary: Developmental Constraints on Learning Artificial Grammars with Fixed, Flexible, and Free Word Order

Aniello De Santo*

Department of Linguistics, Stony Brook University, Stony Brook, NY, United States

Keywords: language learning, cognitive development, artificial grammars, formal language theory, complexity

A commentary on

Developmental Constraints on Learning Artificial Grammars with Fixed, Flexible and Free Word Order

by Nowak, I., and Baggio, G. (2017). *Front. Psychol.* 8:1816. doi: 10.3389/fpsyg.2017.01816

OPEN ACCESS

Edited by:

Stefano F. Cappa,
Istituto Universitario di Studi Superiori
di Pavia (IUSS), Italy

Reviewed by:

Simona Mancini,
Basque Center on Cognition, Brain
and Language, Spain
Cristiano Chesi,
Istituto Universitario di Studi Superiori
di Pavia (IUSS), Italy

***Correspondence:**

Aniello De Santo
aniello.desanto@stonybrook.edu

Specialty section:

This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Psychology

Received: 12 December 2017

Accepted: 19 February 2018

Published: 06 March 2018

Citation:

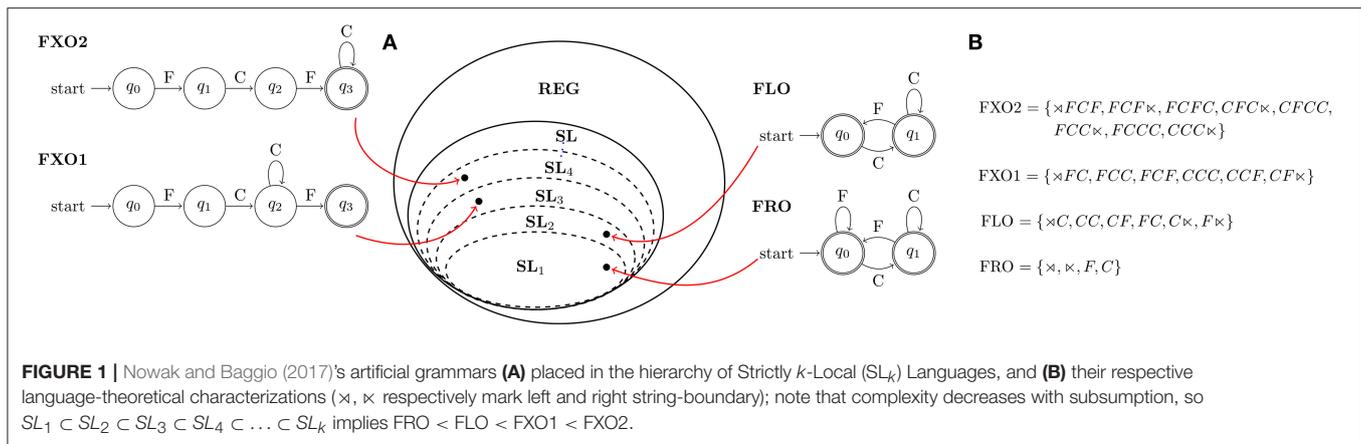
De Santo A (2018) Commentary:
Developmental Constraints on
Learning Artificial Grammars with
Fixed, Flexible, and Free Word Order.
Front. Psychol. 9:276.
doi: 10.3389/fpsyg.2018.00276

A long standing hypothesis in linguistics is that typological generalizations can shed light on the nature of the cognitive constraints underlying language processing and acquisition. In this perspective, Nowak and Baggio (2017) address the question of whether human learning mechanisms are constrained in ways that reflect typologically attested (*possible*) or unattested (*impossible*) linguistic patterns (Moro et al., 2001; Moro, 2016).

Here, I show that the contrasts in Nowak and Baggio (2017) can be explained by language-theoretical characterizations of the stimuli, in line with a relatively recent research program focused on studying phonological generalizations from a mathematical perspective (Heinz, 2011a,b). The fundamental insight is that linguistic regularities that fall outside of certain complexity classes cannot be learned, due to computational properties reflecting implicit cognitive biases.

DEVELOPMENTAL CONSTRAINTS ON LEARNING

In order to test whether adults and children have different biases toward typologically plausible patterns, Nowak and Baggio (2017) construct 4 finite state grammars imposing varying constraints on word-order (fixed: FXO1 and FXO2; flexible: FLO; and free: FRO), instantiated over two word-classes: shorter, more frequent words (F-word) or longer, less frequent ones (C-words). Participants were asked to differentiate between strings produced by the grammar they had been trained on, and strings produced by a different grammar (e.g., FXO1 vs. FLO). Adults succeeded in recognizing fixed and flexible word-order strings (Experiment 1: FXO1 vs. FLO) and failed in recognizing free word-order strings (Experiment 2: FXO2 vs. FRO). In contrast, children could recognize flexible word-order and free word-order strings, but not fixed word-order strings (Experiment 3 and 4, replicating the contrasts of Experiment 1 and 2). The authors attribute these results to the inability of children to acquire *typologically implausible* grammars, suggesting that adults either have distinct constraints on language learning, or are able to employ more general learning strategies.



SUBREGULAR COMPLEXITY

Nowak and Baggio (2017) control for information-theoretical differences (e.g., Shannon entropy; Shannon, 1948) among strings to explicitly refute computational explanations of their results. Crucially, a different computational measure—based on language-theoretical characterizations sensitive to structural properties of the grammars—is dismissed by assuming that the finite-state grammars generating the stimuli lead to languages of equivalent complexity (i.e., regular languages).

This latter assumption is grounded in the Chomsky Hierarchy (Chomsky, 1956), which divides languages (*string-sets*) into nested regions of complexity (*classes*) based on the expressivity of the grammars generating them. However, while regular languages were originally treated as a monolithic unit, it has been shown that they can be decomposed into a finer-grained hierarchy of languages of decreasing complexity—the *Subregular Hierarchy* (McNaughton and Papert, 1971; Rogers et al., 2010). A case has been made for the relevance of this classification for cognition (Rogers and Pullum, 2011; Heinz and Idsardi, 2013; Rogers et al., 2013). Recently, it was posited that the complexity of human language patterns is bound by classes in this hierarchy (the *Subregular Hypothesis*; Heinz, 2010; McMullin, 2016; Graf, 2017), which have been shown to make valuable generalizations across different domains (Aksënova et al., 2016; Aksënova and De Santo, 2017). It also appears that the simpler classes in the hierarchy are more easily learnable by humans (Hwangbo, 2015; Lai, 2015; Avcu, 2017).

Here, my focus is on *Strictly k -Local* (SL_k) languages, which define strings in terms of finite sets of allowed k -grams—contiguous sequences of symbols of length k . Consider *CFCFC* and *CFCFCC*, two well-formed strings for FLO. A strictly k -local grammar is constructed by listing the smallest set of k -grams needed to distinguish between well-formed and ill-formed strings (e.g., **FCFCFC**, *CFCFF*):

$$FLO := \{*\times C, CC, CF, FC, C\times, F\times\}^1.$$

¹ \times , \times mark left and right string-boundary.

Language complexity is measured not by the size of the grammar, but by the minimal length (k) of the substrings needed to generate *all (and only)* its well-formed strings. Thus, FLO is a Strictly 2-Local (SL_2) language. Similarly, FRO is SL_1 , FXO1 is SL_3 , and FXO2 is SL_4 (cf. **Figure 1**). Importantly, SL languages form a proper hierarchy in k : FRO is then the simplest language, while FXO2 is the most complex.

We can now interpret the learnability differences shown for adults vs. children, in light of the subregular complexity of the target string-sets. The contrast between FXO1 and FLO (Experiment 1 and 3) shows that SL grammars are equivalently easy for adults independently of the dimension of the k -grams; while children seem unable to correctly generalize over grammars with complexity greater than SL_2 . Language-theoretical considerations also allow for a deeper understanding of the contrast between FXO2 and FRO (Experiment 2 and 4). In Experiment 2, adults perform well when trained over FXO2: if adults can easily learn SL grammars of any size, this is not an unexpected result. What should come as a surprise is the low performance on FRO, the simplest SL_1 grammar. However, consider that by construction FRO allows for any possible combination of symbols from the alphabet. Therefore, the set of strings generated by FXO2 is a proper subset of the set generated by FRO. Low performance of adults trained on FRO is then expected: since strings from FXO2 are also possible strings for FRO, participants will recognize every string as grammatical, and perform worse on the recognition task. Keeping in mind this possible confound, Experiment 4 (low accuracy when trained on FXO2 vs. FRO) suggests that children might be biased in favor of less restrictive and computationally simpler grammars.

CONCLUDING REMARKS

Nowak and Baggio (2017) present an interesting investigation of developmental biases in language learning mechanisms. I argue that a subregular characterization of their stimuli can help interpret learning differences between adults and children, thus suggesting that the *nature* of the observed biases is in fact intrinsically computational. From this perspective, *unlearnable*

patterns would be those requiring computational resources that exceed what is allowed for a specific cognitive subdomain. What emerges is a strong parallel between language-theoretical approaches, and a research program focused on understanding *possible/impossible* patterns in human languages. Thus, as Jäger and Rogers (2012) suggest, closer collaborations between cognitive scientists and formal language theorists would improve the design and interpretation of artificial grammar experiments targeting human language biases.

REFERENCES

- Aksënova, A., and De Santo, A. (2017). "Strict locality in morphological derivations," in *Proceedings of the 53rd Meeting of the Chicago Linguistic Society (CLS53)* (Chicago, IL).
- Aksënova, A., Graf, T., and Moradi, S. (2016). "Morphotactics as tier-based strictly local dependencies," in *Proceedings of the 14th SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology* (Berlin), 121–130.
- Avcu, E. (2017). "Experimental investigation of the subregular hierarchy," in *Proceedings of the 35th West Coast Conference on Formal Linguistics at Calgary* (Calgary, AB).
- Chomsky, N. (1956). Three models for the description of language. *IRE Trans. Inform. Theory* 2, 113–124. doi: 10.1109/TIT.1956.1056813
- Graf, T. (2017). The power of locality domains in phonology. *Phonology* 34, 1–21. doi: 10.1017/S0952675717000197
- Heinz, J. (2010). Learning long-distance phonotactics. *Linguist. Inq.* 42, 623–661. doi: 10.1162/LING_a_00015
- Heinz, J. (2011a). Computational phonology – part 1: foundations. *Lang. Linguist. Comp.* 5, 140–152. doi: 10.1111/j.1749-818X.2011.00269.x
- Heinz, J. (2011b). Computational phonology – part 2: grammars, learning, and the future. *Lang. Linguist. Comp.* 5, 153–168. doi: 10.1111/j.1749-818X.2011.00268.x
- Heinz, J., and Idsardi, W. (2013). What complexity differences reveal about domains in language. *Top. Cogn. Sci.* 5, 111–131. doi: 10.1111/tops.12000
- Hwangbo, H. J. (2015). "Learnability of two vowel harmony patterns with neutral vowels," in *Proceedings of the The third Annual Meeting on Phonology (AMP 2015)*. (Vancouver, BC).
- Jäger, G., and Rogers, J. (2012). Formal language theory: refining the chomsky hierarchy. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 1956–1970. doi: 10.1098/rstb.2012.0077
- Lai, R. (2015). Learnable vs. unlearnable harmony patterns. *Linguist. Inq.* 46, 425–451. doi: 10.1162/LING_a_00188
- McMullin, K. J. (2016). *Tier-based Locality in Long-Distance Phonotactics?: Learnability and Typology*. Ph.D. thesis, University of British Columbia.

AUTHOR CONTRIBUTIONS

AD reviewed the literature, developed the theoretical stance, and wrote the manuscript.

ACKNOWLEDGMENTS

The author would like to thank Alëna Aksënova, John E. Drury, Thomas Graf, and Jon Rawski for helpful remarks.

McNaughton, R., and Papert, S. (1971). *Counter-Free Automata*. Cambridge: MIT Press.

Moro, A. (2016). *Impossible Languages*. Cambridge, MA: MIT Press.

Moro, A., Tettamanti, M., Perani, D., Donati, C., Cappa, S. F., and Fazio, F. (2001). Syntax and the brain: disentangling grammar by selective anomalies. *NeuroImage* 13, 110–118. doi: 10.1006/nimg.2000.0668

Nowak, I., and Baggio, G. (2017). Developmental constraints on learning artificial grammars with fixed, flexible and free word order. *Front. Psychol.* 8:1816. doi: 10.3389/fpsyg.2017.01816

Rogers, J., Heinz, J., Bailey, G., Edlefsen, M., Visscher, M., Wellcome, D., et al. (2010). "On languages piecewise testable in the strict sense," in *Lecture Notes in Artificial Intelligence, vol. 6149*, eds C. Ebert, G. Jäger, and J. Michaelis (Berlin: Springer), 255–265.

Rogers, J., Heinz, J., Fero, M., Hurst, J., Lambert, D., and Wibel, S. (2013). "Cognitive and sub-regular complexity," in *Proceedings of the 17th Conference on Formal Grammar* (Düsseldorf: Springer), 90–108.

Rogers, J., and Pullum, G. K. (2011). Aural pattern recognition experiments and the subregular hierarchy. *J. Logic Lang. Inform.* 20, 329–342. doi: 10.1007/s10849-011-9140-2

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 623–656. doi: 10.1002/j.1538-7305.1948.tb00917.x

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

The reviewer CC and handling Editor declared their shared affiliation.

Copyright © 2018 De Santo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.