Introduction to the special issue on simplicity in grammar learning

Roni Katzir[1], Timoth J. O'Donnell[2], and Ezer Rasin[1]
[1] Tel Aviv University
[2] McGill University

Simplicity has long been central to philosophy of science, at least in the sense that all things being equal, a more parsimonious theory is better than a more complex one. In modern linguistics simplicity has played a particularly prominent role, with explicit discussion in Chomsky 1951, 1965, Halle 1962, Chomsky and Halle 1968, and much subsequent work. The prominence of simplicity in linguistic theory reflects the importance of learning in this domain: children acquiring a language must choose between many different grammars compatible with the input data, and an intriguing possibility is that their choice, perhaps like that of the scientist, is affected by considerations of simplicity.

The present special issue considers the place of simplicity in grammar learning, focusing on recent computational and theoretical linguistic work but very much building on earlier foundations. In addition to discussing the use of simplicity, the papers in this collection touch on some of the challenges involved in turning simplicity from a guiding intuition into a concrete tool. For example, to what extent would such a tool be limited by the observation that simplicity is always stated with respect to a specific frame of reference? Which, if any, of the various notions of simplicity that have been proposed could support successful grammar learning, and would such a notion adequately model how children generalize from the primary linguistic data? Do observed typological generalizations regarding simplicity in linguistic systems arise from general considerations of stability of simple grammars under repeated iterations of learning across generations, or is there (also) a direct pressure for simplicity? Our hope is

that the papers in this issue both advance the understanding of these questions and serve as an invitation to debate them further.

The paper by Chomsky discusses a fundamental observation in treatments of simplicity: that simplicity must always be stated with respect to a concrete frame of reference. This observation highlights an arbitrariness or subjectivity that might be seen as an obstacle to the scientific use of simplicity. A striking insight of early generative grammar, however, was that in the hands of the cognitive scientist this frame-dependence is in fact an asset rather than a liability. In particular, the very dependence on a frame of reference that makes simplicity subjective also makes it possible to reason from typological and acquisitional generalizations to underlying representational frameworks, providing evidence for those frameworks that make the observed generalizations simple. The paper situates several major theoretical developments in generative linguistics – ranging from early work to very recent additions, and including theorizing about the evolution of universal grammar itself – within the context of simplicity-based considerations.

The paper by Rasin, Berger, Lan, Shefi, and Katzir discusses the right notion of simplicity for grammar learning in light of different notions that have been proposed in the literature. In particular, it considers both grammar simplicity, as in the evaluation metric of early generative grammar, and simplicity of describing the data, which is closely related to the Subset Principle and to Maximum Likelihood approaches. The paper concludes that neither notion is adequate on its own but that a notion that combines them in a certain way – as in Solomonoff's theory of induction, Kolmogorov Complexity, and the principle of Minimum Description Length (MDL) – is adequate and could provide the child with a criterion for comparing hypotheses that seems to match linguistic intuitions in various cases. It illustrates the use of this criterion with an implemented MDL learner for phonological rule systems, reporting simulation results on small corpora that present well-known morpho-phonological challenges from the literature.

The paper by Prickett takes a different perspective on simplicity and grammar learning by associating complexity with layers in the hierarchy of formal languages. The paper considers evidence from experiments of artificial grammar learning that suggests a bias for hy-

potheses that reside on lower rungs of the hierarchy over hypotheses that are higher and require greater weak generative power. The paper provides evidence that the implicit learning bias of a specific recurrent neural network is compatible with this kind of preference.

The paper by Johnson, Gao, Smith, Rabagliati, and Culbertson looks at simplicity in morphological paradigms in light of a distinction due to Ackerman and Malouf (2013) between e(numerative) and i(ntegrative) complexity. E-complexity tracks the number of surface distinctions in a paradigm and varies greatly across paradigms. I-complexity is a measure of how predictable the elements of a paradigm are from the form of a representative element and has been argued by Ackerman and Malouf to be consistently low across paradigms. As Ackerman and Malouf note, the typological generalization might be related to learning, though this relation could in principle be indirect and arise from the fact that predictable paradigms might be more stable to intergenerational transmission than unpredictable paradigms. Using both an artificial grammar learning experiment and simulations with RNNs, and using specific information-theoretic formulations of e- and i-complexity, Johnson et al. ask whether there is a more direct learning preference for simpler paradigms. They indeed report such a preference but conclude that it is greater for the typologically variable e-complexity than for the typologically stable i-complexity. They also explore the relation between the two notions of complexity across artificially-generated paradigms and find an inverse relation between the two.

The final paper in this special issue, by Lambert, Rawski, and Heinz, looks at grammar learning through the prism of yet another notion of complexity: that of resource (specifically, space and time) complexity. The paper provides a systematic exploration of representations and learning algorithms that vary in terms of their resource complexity, drawing a connection between the possible combinations of representations and algorithms on the one hand and the subregular hierarchy in phonological typology (Heinz and subsequent work) on the other hand.

REFERENCES

Farrell ACKERMAN and Robert MALOUF (2013), Morphological Organization: The Low Conditional Entropy Conjecture, 89(3):429–464.

Noam CHOMSKY (1951), Morphophonemics of Modern Hebrew, Master's thesis, University of Pennsylvania.

Noam CHOMSKY (1965), Aspects of the Theory of Syntax, MIT Press, Cambridge, MA.

Noam CHOMSKY and Morris HALLE (1968), The Sound Pattern of English, Harper and Row Publishers, New York.

Morris HALLE (1962), Phonology in generative grammar, Word, 18(1/2):54–72.

Roni Katzir
(iD) 0000-0002-0241-1896
rkatzir@tauex.tau.ac.il

Department of Linguistics
and Sagol School of Neuroscience
Tel Aviv University
Tel Aviv, Israel 6997801

Ezer Rasin
(iD) 0000-0001-8980-5566
rasin@tauex.tau.ac.il

Department of Linguistics
Tel Aviv University
Tel Aviv, Israel 6997801

Timothy J. O'Donnell
(iD) 0000-0000-0000-0000
timothy.odonnell@mcgill.ca

Department of Linguistics
McGill University
Montréal, Canada