

More learnable than thou? **Testing metrical phonology representations with child-directed speech**

One (often implicit) motivation for a knowledge representation (**KR**) comes from an *argument from acquisition* [13], with the idea that language acquisition is straightforward if children's hypothesis space is defined by the correct KR (e.g., [1][3][4]). Acquisition is then the process of selecting the correct grammar from that hypothesis space, given language input. Based on an argument from acquisition, we establish quantitative metrics for comparing KR's that assess learnability from realistic acquisition data. We conduct a computational learnability analysis [10][13] for three KR's proposed for metrical phonology and test them on English, a language that is notoriously noisy with respect to metrical phonology. We find that all three KR's have similar learnability potential, but the proposed English grammars within each KR vary on the amount of English child input data they can account for. Notably, the English grammar in all three KR's is *not* the grammar able to account for the most English child input data, even if the learner has some knowledge of the interaction between metrical phonology and morphology. This suggests learnability issues exist for the proposed English grammar, no matter what the KR, given a learner attempting to learn a generative system that accounts for as much of the input data as possible.

We investigate two parametric KR's [5][8] and a constraint-based KR [6][12] by establishing and evaluating several formally defined learnability-based metrics: (i) a grammar's *compatibility* with ~4800 types (~100,000 tokens) of American English child-directed speech data (Brent corpus: CHILDES [11]), (ii) each KR's *learnability potential*, given all the grammars it defines, and (iii) the *relative compatibility* of the KR's proposed English grammar, compared against other grammars defined by the KR. English is an excellent test case for competing KR's because of its difficulty: there are numerous exceptions, so no one grammar (no matter what the KR) is compatible with all the data [13]. Moreover, there are known interactions between English metrical phonology, morphology [2][7][9], and grammatical category [6][7]. We consider both purely phonological instantiations of each KR and instantiations that include some morphological knowledge.

We find that all three KR's have similar learnability potential. In contrast, there is variation in the compatibility of each KR's proposed English grammar, with one of the parametric English grammars performing the best. Nonetheless, all three KR's suffer from a similar problem: the English grammar's compatibility is lower than a significant number of other grammars defined by the KR. This means an unbiased learner looking for the optimal grammar that accounts for the observed data would not choose the English grammar in any of these KR's when given English child-directed input. Interestingly, we also find that knowing English inflectional morphology is stressless does not significantly aid data compatibility, and so would not aid acquisition.

We discuss which aspects of the proposed English grammars may be hurting learnability, observing that fairly small changes in parameter values or constraint-rankings may lead to significantly higher compatibility. We additionally discuss ways a learner may still be able to learn the English grammar from English input by incorporating (i) additional linguistic knowledge about morphology and grammatical category and (ii) learning biases, such as accounting only for the stress data viewed as productive [10].

REFERENCES

- [1] Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- [2] Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.
- [3] Crain, Stephen & Paul Pietroski. 2002. Why language acquisition is a snap. *The Linguistic Review* 19. 163-183.
- [4] Dresher, Elan. 1999. Charting the learning path: Cues to parameter setting. *Linguistic Inquiry* 30. 27-67.
- [5] Halle, Morris & Jean-Roger Vergnaud. 1987. *An Essay on Stress*. Cambridge, MA: MIT Press.
- [6] Hammond, Michael. 1999. *The Phonology of English*. Oxford: Oxford University Press.
- [7] Hayes, Bruce. 1982. Extrametricality and English stress, *Linguistic Inquiry*, 13, 227-276.
- [8] Hayes, Bruce. 1995. *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.
- [9] Kiparsky, Paul. 1979. Metrical Structure Assignment is Cyclic. *Linguistic Inquiry*, 10.4, 421-441.
- [10] Legate, Julie & Yang, Charles. 2012. Assessing child and adult grammar. In Berwick, R. & Piattelli-Palmarini, M. (Eds.), *Rich Languages from Poor Inputs*. Oxford: Oxford University Press.
- [11] MacWhinney, Brian. 2000. *The CHILDES Project: Tools for analyzing talk*. Mahwah, NJ: Lawrence Erlbaum Associates.
- [12] Pater, Joe. 2000. Non-uniformity in English secondary stress: the role of ranked and lexically specific constraints. *Phonology*, 17, 237-274.
- [13] Pearl, Lisa. 2011. When unbiased probabilistic learning is not enough: Acquiring a parametric system of metrical phonology. *Language Acquisition*, 18(2), 87-120.