Learners change artificial languages to constrain free variation in line with typological principles

Shannon Mooney and Youngah Do (Georgetown University and University of Hong Kong)

A substantive bias in adult acquisition of phonological patterns has been argued (Wilson 2006), though experimental work has yielded mixed results (Moreton and Pater 2012). Adults, in contrast to children, also have been shown to simply reproduce the frequencies of morphosyntactic variants in their input in an experimental paradigm, the learning result of which barely shows any role of biases (Hudson Kam and Newport 2005 and 2009, though see also recent work by Smith et al. and Samara et al. which suggests initially-unnoticeable biases that magnify over generations of transmission). Our study attempts to induce free phonological variation in an artificial language to adult learners to test the role of a substantive bias. We take the typology for rounding harmony asserted by Kaun (2004) as a test case for substantive bias effects and test if our participants will shift towards variant distributions that mimic typological patterns, in contrast to the results from studies of adult acquisition of morphosyntactic variation.

Kaun (2004) describes typological asymmetries of rounding harmony that favor (1) a non-high trigger vowel, (2) a high target vowel, and (3) height agreement between trigger and target vowels. She also describes cases of free variation in rounding harmony attested in Tuha, Tofa, and Altai Tuvan (Turkic). In Tuha and Tofa, the target is always high, and, in Tofa, rounding harmony is obligatory when there is height agreement. In our study, an experimental component exposed learners to a language characterized by variable rounding harmony. The language exhibited high and non-high round triggers, and high and non-high targets that underwent rounding harmony at an identical frequency across all contexts. In one artificial language, the frequency of rounding harmony extending from trigger to target was 66%, and in the other language the frequency of rounding harmony was 33%. In all other ways, the two languages were equal and balanced for vowel distributions.

Learners heard in their input multiple tokens of a single noun, some tokens with rounding harmony triggered to the affix, and some without. English-speaking undergraduate student participants (n=32) underwent a 20-minute training, hearing sentences paired with scenes. Our language consisted of 5 nouns, 2 of which were only introduced during the testing phase, and 2 verbal particles. Half the participants were exposed to the language exhibiting variable rounding harmony across the final syllable of the noun and the following single syllable verbal particle at the higher frequency (66%). The others received the lower frequency of rounding harmony (33%). In both languages, we ensured equal probability of rounding harmony in all lexical items and both non-high and high vowels. Rounding harmony was in true free variation rather than phonological alternation; the way the language was set up, a noun was paired with a following verbal affix that optionally experienced rounding harmony in accordance with the final vowel of the noun to which it attached. Testing consisted of exposure to novel scenes with novel lexical items, with participants recorded while describing the scenes in the language. This resulted in 96 utterances collected per participant, where rounding harmony would be productively but variably applied to novel lexical items.

Learners boosted more "natural" rounding harmony patterns and significantly reduced rounding harmony in "unnatural" contexts relative to their input, in line with some of the typological principles for rounding harmony proposed by Kaun (2004). Input variation came to be constrained by vowel height in participant production. Specifically, despite equal input distribution of rounding harmony across high and non-high vowels, and agreeing vs.

non-agreeing triggers and targets, a linear regression shows participants had significantly higher rates of rounding harmony in high targets with height agreement ($p=.02$) and without height agreement ($p=.03$) compared to non-high targets. This suggests to us that a substantive bias may operate to shift phonological knowledge of the artificial language towards more natural patterns, resulting in a distribution consistent with natural language rounding harmony principles. Despite any operable substantive bias, participants maintained conditioned probabilistic variation in their productions rather than introducing any categorical patterns to alternations.

We then simulated learning of our experimental results using MaxEnt modeling. Training data was the two constructed variable rounding harmony languages (66% and 33% rounding harmony with free variation) as reproduced by participants. We hypothesized that the trained grammars would be illustrative of a participant substantive bias.

Simulations of both the more and less rounding harmony languages favor non-high triggers and/or high targets for rounding harmony over high triggers and/or non-high targets. Surprisingly, however, the reported typological principle that height agreement between trigger and target promotes rounding harmony is not reflected in the learned grammar of the current simulation. Given the structural simplicity of the artificial languages tested (i.e., N+V structure with systematic local rounding harmony), we go so far as to advocate that the constraints uncovered by this experiment are true phonetically-grounded inductive biases, while preference for height agreement in rounding harmony, not found in this experiment, may be a structural bias towards simplicity and against new information (similar to the sort found across domains of cognition), and, if so, would not be as robust in the learning of structurally-simple languages.

A newly uncovered (to our knowledge) substantive bias comes to weakly constrain phonological variation in our structurally-simple artificial languages. The emergence of this apparent bias bridges a longstanding difference in results between artificial language paradigms used in the morphosyntactic and phonological research traditions, leading to new directions for experimental studies of the role of substantive bias in phonological acquisition.

Works Cited

Hudson Kam, Carla L., and Elissa L. Newport. (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., and Elissa L Newport. (2009). Getting It Right by Getting It Wrong: When Learners Change Languages. *Cognitive Psychology* 59(1): 30–66.

Kaun, Abigail. (2004). The typology of rounding harmony. *Phonetically based phonology*, 87416.

Moreton, Elliott, and Joe Pater. (2012). Structure and Substance in Artificial-Phonology Learning, Part II: Substance. *Language and Linguistics Compass* 6(11): 702–718.

Samara, Anna, Kenny Smith, Helen Brown, and Elizabeth Wonnacott. (2017). Acquiring Variation in an Artificial Language: Children and Adults Are Sensitive to Socially Conditioned Linguistic Variation. *Cognitive Psychology* 94: 85–114.

Smith, Kenny, Amy Perfors, Olga Fehér, et al. (2017). Language Learning, Language Use and the Evolution of Linguistic Variation. Philosophical Transactions of the Royal Society B: Biological Sciences 372(1711): 20160051.

Wilson, C. (2006). Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive science*, 30(5), 945-982.