A Bayesian meta-analysis of inductive bias in phonological learning

Canaan Breiss[1] (cbreiss@usc.edu), Anna Runova[1] (runova@usc.edu)

[1]University of Southern California

Background: A core question of cognitive science is the nature of inductive biases that guide language acquisition, allowing infants to learn efficiently, and acting as a possible source for typological asymmetries [1]. In phonological learning, researchers have posited a *simplicity bias* (a preference for structurally simpler patterns – putatively domain-general) and a *naturalness bias* (a preference for more phonetically natural patterns – putatively language-specific) (cf. [2],[3]). Over the last 20 years or so, Artificial Grammar Learning (AGL) experiments have become a widely-used technique to empirically investigate the nature of these learning biases. In these studies, adults are taught, and then generalize, a toy language with specific properties, and researchers seek evidence for an advantage in learning speed or accuracy of some generalizations over others. However, it remains difficult to draw large-scale firm conclusions from this literature (cf. [4], [5], [6]), due to differences in the definition and operationalization of core concepts (e.g. “simplicity” or “naturalness”), diverse experimental methodologies that tap into different parts of the larger question (iterated learning, training with feedback, poverty of the stimulus, etc.), small sample sizes, and a lack of replication studies. Instead, the current state of the literature reflects many independent small-scale AGL investigations of possible inductive biases suggested by observed typological asymmetries, which although useful, are difficult to integrate into a cohesive understanding of inductive biases in phonology at the scale of the whole learner. Here we seek to provide a summatory, integrative perspective that is crucial if we are to have a detailed, mechanistic, and implemented model of relevant human cognitive processes, including inductive bias, that guide language acquisition and shape typology (cf. [7], [8]).

Methods: We use a Bayesian meta-analysis to examine the results of AGL studies of naturalness bias, focusing primarily on vowel harmony and other vowel-related phonological patterns (work on consonants in in progress). Following the PRISMA guidelines for meta-analysis [9], we screened papers from the relevant literature based on our inclusion guidelines (i.e., studies of neurotypical adults’ behavioral dependent variables, examining generalization of an artificial language), and extracted study characteristics and statistical results. This yielded 21 papers containing data from 34 experiments, resulting in 97 effect sizes, reflecting data from 1,416 participants. We divided papers into three groups: those that compared learnability of natural patterns to chance (fig. 1), unnatural patterns to chance (fig. 2), and natural to unnatural (fig. 3). We analyzed standardized effect size (Hedge’s *g*) using a Bayesian mixed-effects meta-analytic regression model in Stan [10], integrating the three measurement types with a custom likelihood.

Results: We find strong evidence that patterns deemed “natural” by the authors of the paper have a larger effect size (Table 1, row 2). What “natural” means, however, is unclear given the heterogeneity of definitions. Turning to patterns which are argued by the authors to be natural because they have *phonetic precursors* (vowel harmony patterns, nasality agreement, place assimilation in consonants), we find very little evidence that they are learned better (Table 1, row 3). Another operationalization of phonetic naturalness, the *number of changing phonological features* that are involved in an alternation, is a fairly strong predictor of an increased effect size (Table 1, row 5), mirroring qualitative summaries of the literature ([2],[3]), though uncertainty is still quite high. However, patterns with larger numbers of changing features may also be more acoustically distinct, thus confounding formal (feature-based) and functional (perceptual) explanations. Moreover, we find that almost all included studies use auditory stimuli at training and test, and employ a forced-choice design, making it impossible to disentangle the role of number of changing features and acoustic distinctiveness. Finally, on the methodological front, we find that there is no evidence that AGL experiments carried out online differ meaningfully in effect size from those carried out in-lab (Table 1, row 6), supporting the validity of this increasingly-common design choice.

In general, we find that the large amount of heterogeneity in the literature leads our model to indicate that any individual study is likely overconfident about the certainty of its results. Further, the large imbalance in the number of studies with different manipulations makes our model quite uncertain in its estimates (second and third columns in Table 1 indicate the number of effect sizes on each side of a comparison).

References

[1] Chomsky 1965. Aspects of the theory of syntax. MIT Press.

[2] Moreton, E., & Pater, J. (2012a). Structure and Substance in Artificial-phonology Learning, Part I: Structure. Language and Linguistics Compass, 6(11), 686–701

[3] Moreton, E., & Pater, J. (2012b). Structure and Substance in Artificial-Phonology Learning, Part II: Substance. Language and Linguistics Compass, 6(11), 702–718. https://doi.org/10.1002/lnc3.366

[4] Glewwe, E. R. (2019). Bias in phonotactic learning: Experimental studies of phonotactic implicationals [PhD Thesis, University of California, Los Angeles].

[5] Do, Youngah, and Jonathan Havenhill. "Production and substantive bias in phonological learning." Proceedings of the Annual Meetings on Phonology. Vol. 8. 2021.

[6] Zheng, S., & Do, Y. (2025). Substantive Bias in Artificial Phonology Learning. *Language and Linguistics Compass*, *19*(1), e70005.

[7] Kiparsky 1965. *Phonological Change*. Doctoral dissertation, MIT.

[8] White, James. "Evidence for a learning bias against saltatory phonological alternations." *Cognition* 130.1 (2014): 96-115.

[9] Page M J, McKenzie J E, Bossuyt P M, Boutron I, Hoffmann T C, Mulrow C D et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews BMJ 2021; 372

[10] Stan Dev. Team, 2025 Stan Reference Manual, v. 2.36. https://mc-stan.org
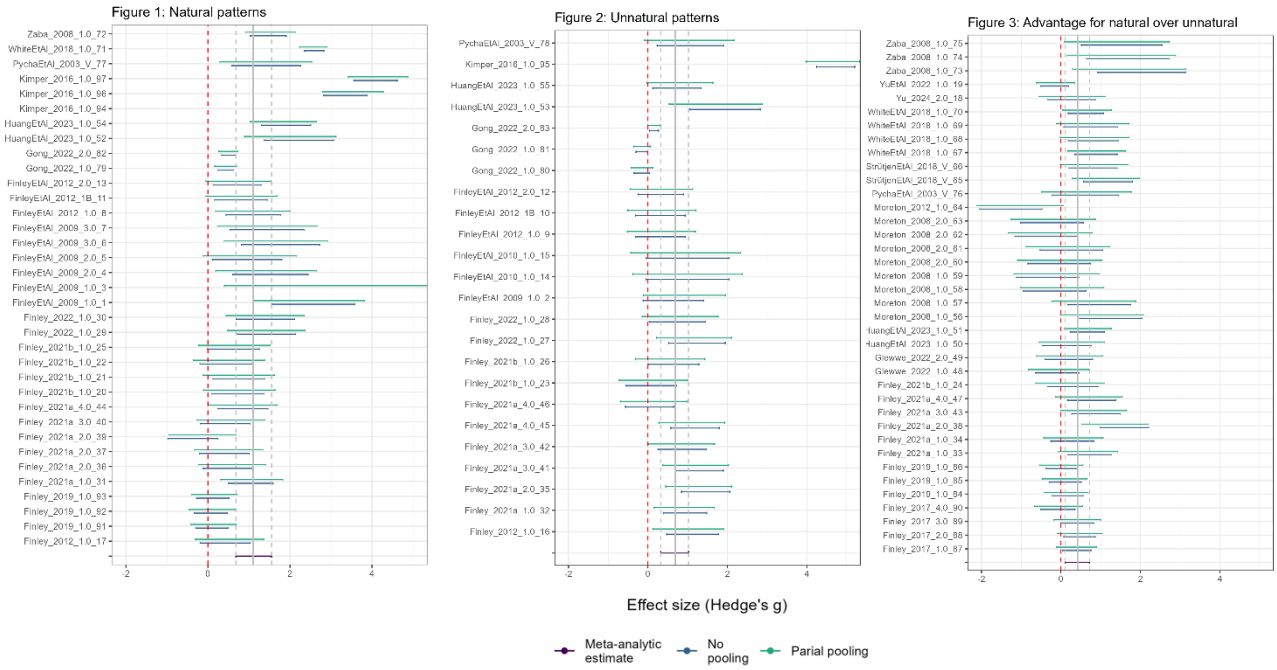
Table 1 (results of LMM integrating across three comparison datasets)	*n* positive	*n* negative	Posterior median effect size, w/ 95% Cred. Int.	Prob. non-zero effect
"Natural" vs. "unnatural"	54 natural	43 unnatural	.42 [.11, .72]	.99
Phonetic precursors	87 yes	10 no	.07 [-.25, .40]	.67
Featural agreement	85 yes	12 no	.013 [-.18, .44]	.81
Num. features changing	2 features = 28	1 feature = 65	.17 [-0.16, .50]	.84
Location	83 in-lab	14 online	0.01 [-.43, .45]	.51