Complexity and naturalness biases in phonotactics: Hayes and White (2013) revisited

Brandon Prickett - University of Massachusetts, Amherst

Annual Meeting on Phonology 2017

Introduction

- Hayes and White (2013) found evidence that English speakers learn **natural** phonotactic patterns in their native language better than unnatural ones.
  - ⇒ Patterns, in the form of constraints, were found by the UCLA Phonotactic Learner (Hayes and Wilson 2008).
  - ⇒ Constraints were considered natural if they made words **"easier to articulate"** or more **"perceptually distinct"** (Hayes and White 2013:47).
- However, they did not control for the **complexity** of the constraints they tested.
  - ⇒ Complexity has been shown to affect laboratory learning of phonology, as well as typological trends (see Moreton and Pater 2012 for a review).
- In this study, I used Hayes and White's (2013) methodology, but added complexity as an additional variable of interest when classifying constraints.
  - ⇒ I used **feature counting** as a metric for constraint complexity, following Chomsky and Halle (1968).
  - ⇒ For other complexity metrics in phonology, see Heinz et al. (2009), Rasin and Katzir (2016), and Moreton et al. (2017).

- I chose 15 constraints found in English by the UCLA Phonotactic Learner (Hayes and Wilson 2008):

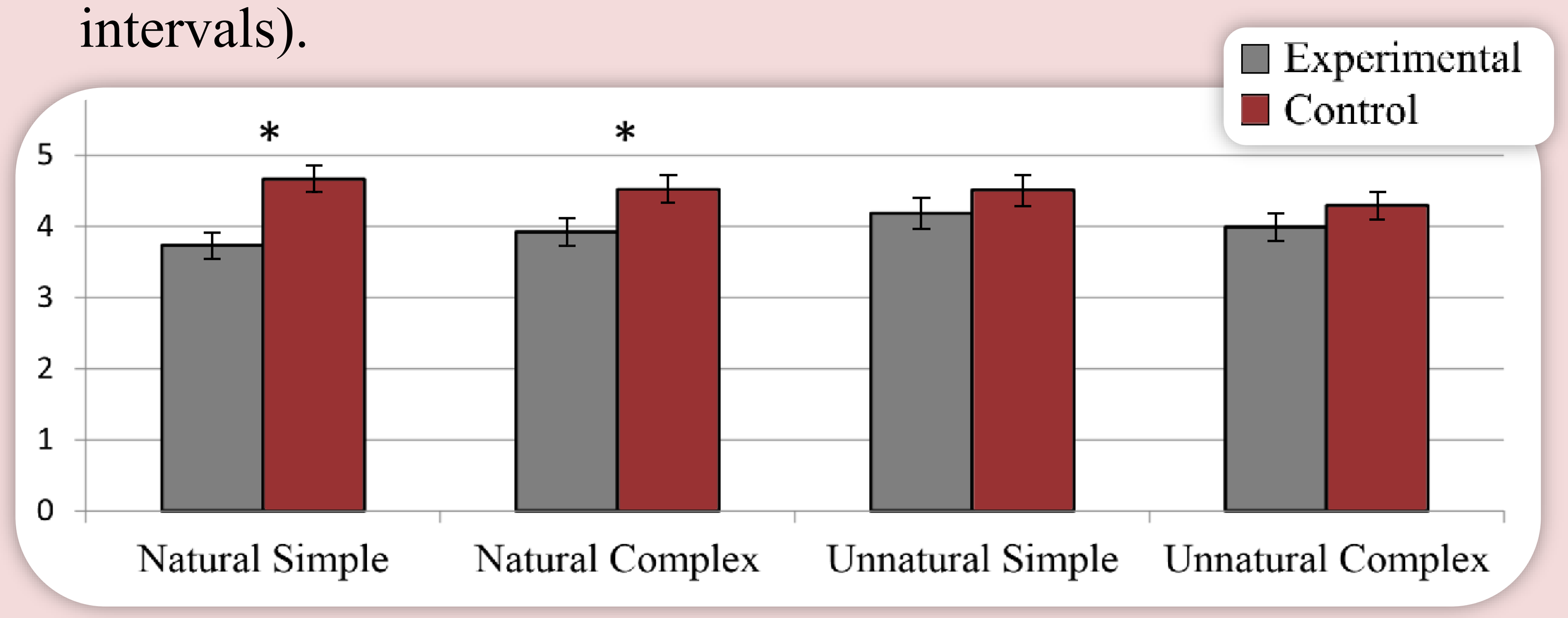
Constraints	Descriptions
Natural Simple	
*[+syllabic][+syllabic]	No hiatus
*[-cons.][+cons.] in onset	Sonority ordering in onsets
*[-cont.][-cont.] in onset	No sonority plateau in onsets
*[-son.][+son] in coda	Sonority ordering in codas
Natural Complex	
*[+cons.,-anterior][-son.]	No pre-obstruent palatals
*[+diph., +round][+labial]	No round diphs before labial cons
*[-low, +tense][+son., +dors.]	No [-low] vowels before low cons
*[+nas., +cor.][+dors.] in coda	No heterorganic nasal clusters in coda
Unnatural Simple	
*[+diph.][+cont.,-anterior]	No diph. before palatal fricatives
*[+round,-back][-anterior]	No [oɪ] before palatals or [ɹ]
*[-son.][-low] in coda	No obstruents in a coda before glides
Unnatural Complex	
*[+round, +high][-cons.,-son.]	No [u, ʊ] before [h]
*[+diph., +round][+vce., +dors.]	No round diphthongs before [g]
*[+cons.,-ant.][+high,-syllabic]	No palatals before glides
*[+diph., +round.][-strid.] in onset	No round diphthongs before non-stridents, except when the non-strident is in coda position

Procedure

- Participants (N=77) were presented with nonce words via audio recordings and orthographic representations and were asked to rate how "good" each stimulus sounded as a word in English.
  - ⇒ Ratings were made using Magnitude Estimation (Stevens 1975).
  - ⇒ As in Hayes and White (2013), *poik* was given as a baseline with a rating of 100.
- There were two types of stimuli:
  - ⇒ **Experimental**: these stimuli each violated one of the constraints of interest while violating no other phonotactic constraints in English.
  - ⇒ **Control**: each of these stimuli had an experimental partner. They differed from their partner in only one segment and violated no English constraints.

Results

- The figure below shows the mean log ratings given to each category of stimulus (error bars represent 95% confidence intervals).



- I subtracted the log rating for each violating stimulus from its non-violating partner, ran a repeated measures ANOVA (with an error term for subjects) on these data, and found that naturalness, complexity, and their interaction were significant:

	Mean Sq	F value	Pr(>F)
Natural	78418	176.76	< 2e-16 ***
Simple	18936	42.68	4.15e-10 ***
Natural:Simple	3647	8.22	0.00453 **

(For residuals: df=228, Sum Sq=101150, and Mean Sq=444)

Issues

- Problems with complexity in this experiment:
  - ⇒ Constraint complexity is highly dependent on which feature theory is used. See Hayes and White (2013) for more on the features used here.
  - ⇒ It's also unclear how prosodic features, like syllabic information, should be treated by a feature-counting metric (I counted them as half a feature each).
- Why does naturalness have such a strong effect here when it hasn't been shown to affect artificial language learning? (see Moreton and Pater 2012 for a review)
  - ⇒ This could be the result of differences between natural and artificial language learning (i.e. in real-world scenarios, learning has a naturalness bias, but in laboratory settings it doesn't).
    - ◊ There's some evidence to suggest that a more realistic learning setting could elicit more of a naturalness bias (Peperkamp and Martin 2016).
  - ⇒ Gaps in the UCLA Learner's training data (e.g. novel words) could be more likely to violate unnatural constraints. For example:
    - ◊ ***[+diph., +round][+vce., +dors.]** violated by "Smaug" (*The Hobbit*)
    - ◊ ***[+cons.,-ant.][+high,-syll.]** violated by "schwifty" (*Rick and Morty*)
- Other issues also arise from the Hayes and White (2013) methodology:
  - ⇒ Stimuli need to be presented orthographically, since ungrammatical words can sometimes be misperceived. However, this also means that subjects' ratings could be ratings of **orthographic well-formedness**.
  - ⇒ **Compromises** must be made in regards to the constraints that are tested, since constraints are chosen from a fairly small set (~160).
    - ◊ E.g. Hayes and White's (2013) complexity confound,
    - ◊ My small number of unnatural simple constraints,
    - ◊ The small difference between my complex and simple categories,
    - ◊ Etc...

Conclusion

- This study **successfully replicated Hayes and White's (2013) finding** that naturalness has an effect on which phonotactic generalizations speakers learn.
- The effect of complexity and the interaction between the two were also significant, suggesting that **both variables play a role in phonological acquisition**.

References

Chomsky, Noam and Morris Halle (1968). *The Sound Pattern of English*. Harper & Row. Hayes, Bruce, and James White (2013). Phonological Naturalness and Phonotactic Learning. *Linguistic Inquiry 44.1*: 45-75.Hayes, Bruce, and Colin Wilson (2008). A Maximum Entropy Model of Phonotactics and Phonotactic Learning. *Linguistic Inquiry 39.3:* 379-440. Heinz, Jeffrey, Gregory Kobele, and Jason Riggle (2009). Evaluating the Complexity of Optimality Theory. *Linguistic Inquiry, 40(2)*: 277-288. Moreton, Elliott, and Joe Pater (2012). Structure and substance in artificial-phonology learning. Part I, Structure. Part II, Substance. *Language and Linguistics Compass 6 (11)*: 686-701 and 702-718. Moreton, Elliott, Joe Pater, and Katya Pertsova (2017). Phonological concept learning. *Cognitive science*, *41(1)*: 4-69. Peperkamp, Sharon and Alexander Martin (2016). Sleep-dependent consolidation in the learning of natural vs. unnatural phonological rules. Presented at *The 15th Conference on Laboratory Phonology*. Rasin, Ezer and Roni Katzir (2016). On Evaluation Metrics in Optimality Theory. *Linguistic Inquiry 47:2*: 235-282. Stevens, Stanley S. (1975). Psychophysics: Introduction to its Perceptual, Neural, and Social Prospects. New York: John Wiley.