Not all wug-tests are created equal: Cognitive load impairs access to the phonological grammar

Claire Moore-Cantwell & Dave Kush, UCLA & NTNU

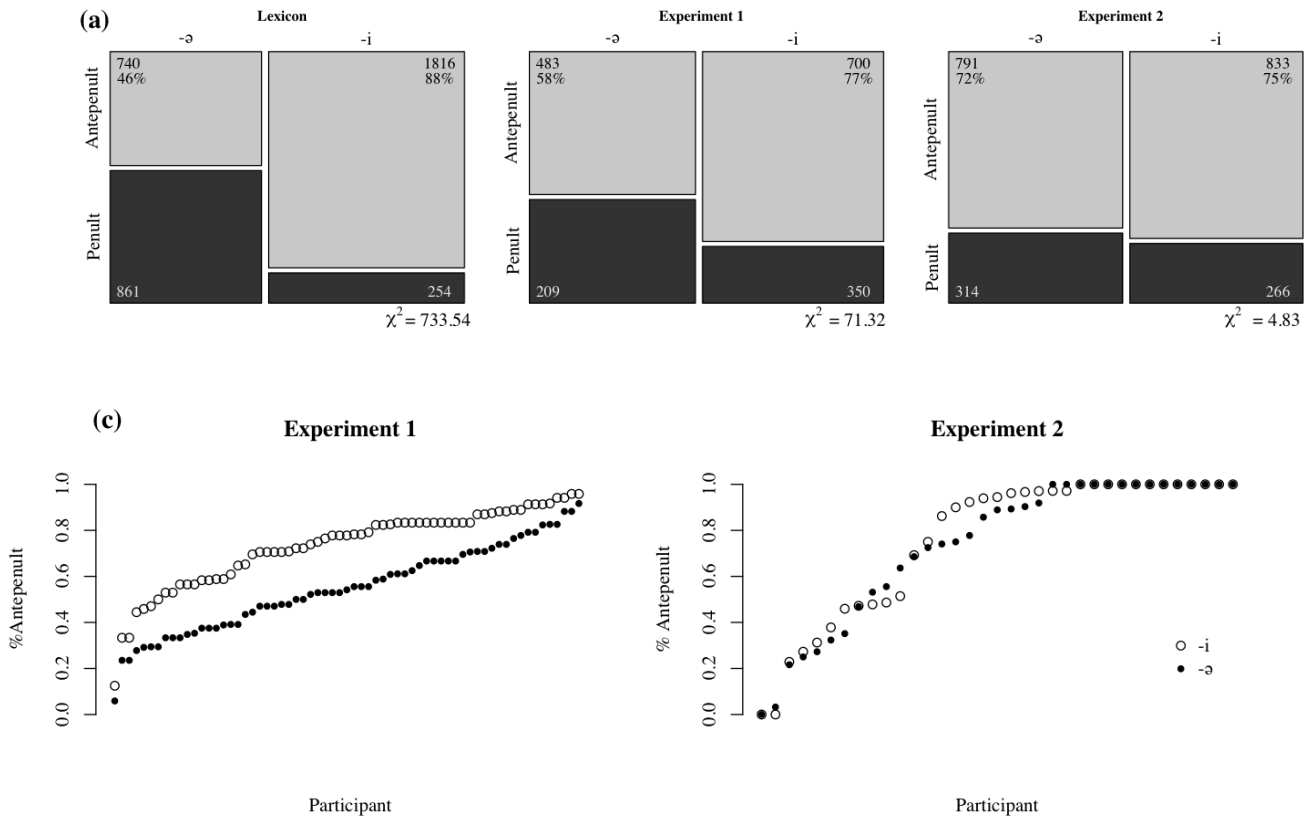
When a variable phonological pattern exists in the lexicon of a language, speakers exhibit the ability to 'probability match' on novel words, producing a distribution of output forms which matches the distribution of form types found in the lexicon, even following quite complex and subtle statistical trends (Ernestus and Baayen, 2003; Hayes et al., 2009 et. seq). This result has led to the adoption of models of the Phonological Grammar which are inherently probabilistic (Goldwater and Johnson, 2003; Hayes and Wilson, 2008). In this paper we argue that increased cognitive load during an experiment can impair speakers access to this complex, probabilistic grammar. Exp. 1 shows 'probability matching' behavior, and consistency across participants, while Exp. 2 increases participants cognitive load with a memorization task, and shows less probability matching, more categorical behaviour from each participant, and inconsistency across participants.

The probabilistic trend: In English words longer than two syllables, stress is typically penultimate ('banána') or antepenultimate ('Cánada'). A search of the CMU pronouncing dictionary (Weide, 1994) revealed that [i]-final words were biased towards taking antepenultimate stress, and [ə]-final words were unbiased. In words at least 3 syllables long, 88% of i-final words were antepenultimately stressed, but only 54% of ə-final words, were antepenultimately stressed (a).

Methods: Nonwords (half -i, half -ə) were constructed so as to have very sparse neighborhoods (less than 0.01) according to the Generalized Neighborhood Model (Bailey and Hahn, 2001). Nonwords were presented auditorily as three individual syllables with acoustically ambiguous stress ([bæ] [mæ] [ki]). **Exp. 1:** 104 participants, recruited through Amazon Mechanical Turk, were recorded as they spoke the syllables fluently as a single word. Next, participants 'transcribed' their own production by listening to 2 versions of the nonword ([bǽməki], [bəmǽki]) and selected the version most similar to what they produced. Exp. 2: 39 participants in a lab saw a real English word (10 each -i, antepenult stress, -əantepenult, -i penult, -əpenult) printed on the computer screen, then heard a nonword, then spoke first the real word then the nonword. Stresses were transcribed.

Results: Data from 66 participants in Exp. 1 was analyzed, all at least 90% accurate in their 'transcriptions'. Participants extended the probabilistic trend in the lexicon to nonwords (a). In Exp. 2, participants reported difficulty with the task, and produced many errors and disfluencies on real words and nonwords. Data was analyzed from 35 participants who produced errors on fewer than half of trials. Overall, the lexical trend was only very weakly extended to nonwords (a). In (b) each participant's rate of choosing antepenult stress is shown for the two studies. In Exp. 1 participants were extremely consistent. Although they varied in their overall preference for antepenult stress, every participant varied in their stress choices and every participant produced more antepenult stress on i-final items. The same was not true in Exp. 2. Many participants always, or nearly always, produced one kind of stress (usu. antepenult). And while some participants produced more antepenult stress on i-final items, many did not.

Discussion: Because participants exhibit greater consistency in Exp. 1, we argue that their behavior there more directly reflects the Phonological Grammar which all English speakers share. In Exp. 2, we argue, the increased cognitive load imposed by the priming task leads participants to ignore certain grammatical constraints during their choice of stress on novel words. The phonological grammar is modeled using a set of many weighted constraints (Goldwater and Johnson, 2003), which allow speakers to make good guesses about a word's pronounciation, whether it is novel or they have forgotten parts of it (Ernestus and Baayen, 2001). Under cognitive load, speakers' access to these constraints is impaired, leading them to select only the highest-weighted ones. Categorical behavior arises when the chosen constraints do not conflict.



References

Todd M. Bailey and Ulrike Hahn. Determinants of wordlikeness: Phonotactics or lexical neighborhoods? *Journal of Memory and Language*, 44:568–591, 2001. doi: 10.1006/jmla.2000.2756.

Mirjam Ernestus and Harald Baayen. Choosing between the Dutch past-tense suffixes *-te* and *-de*. In Ton van der Wouden and Helen de Hoop, editors, *Linguistics in the Netherlands*, pages 81–93. john Benjamins, Amsterdam, 2001.

Mirjam Ernestus and Harald Baayen. Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language*, 79(1):5–38, 2003. doi: 10.1353/lan.2003.0076.

Sharon Goldwater and Mark Johnson. Learning OT constraint rankings using a maximum entropy model. In Jennifer Spenader, Anders Eriksson, and Osten Dahl, editors, *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*, pages 111–120, 2003.

Bruce Hayes and Colin Wilson. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry*, 39:379–440, 2008. doi: 10.1162/ling.2008.39.3.379.

Bruce Hayes, Kie Zuraw, Péter Siptár, and Zsuzsa Londe. Natural and unnatural constraints in Hungarian vowel harmony. *Language*, 85(4):822–863, December 2009. doi: 10.1353/lan.0.0169.