Gradient and categorical constraints: evidence from Cantonese

Classical generative phonology draws a distinction between possible words (*accidental gaps*) and impossible words (*systematic gaps*) [4]. This model predicts that speakers should categorically judge accidental gaps to be well-formed and systematic gaps to be ill-formed, and thus rate accidental gaps more highly than systematic gaps in a wordlikeness judgment task. Such *structure models* attribute the listener's differing judgments to abstract phonological generalizations, such as a ban on e.g. [coronal][coronal] sequences [5]. However, recent research into well-formedness suggests speaker judgments of both words and nonwords to be gradient in nature, and influenced by lexical as well as phonotactic factors [1] [2] [3] [6] [7]. This class of models, sometimes known as *unit models*, attribute gradient judgments to differences in experience, such as unit frequency or lexical density. However, most nonword research has focused on accidental gaps. We designed a wordlikeness judgment task to explore judgments of both systematic and accidental gaps in Cantonese.

Background. Cantonese contains 19 consonants /p p^h t t^h ts ts^h k k^h k^w k^{wh} m n ŋ f s h l j w/, 8 monophthongs /aː a ɛ i ɔ ø u y/, 11 diphthongs /ai ɐi au ɐu ei ɛu ɵy ɔi ui iu ou/ and 6 tones /55 35 33 21 13 22/. The systematic gaps examined here are a labial dissimilation gap (labials, labiovelars, and /w/ do not occur as onsets in syllables with labial codas; labial codas cannot occur with rounded vowels), an onset-tone gap (unaspirated onsets do not occur with 11 or 23 tones, aspirated onsets do not occur with 22), and a coronal-vowel gap (coronal onsets and codas may not occur with the nuclei /o, u/).

Methodology. The corpus was a set of 432 CV(C) words and nonwords, derived eight onset phonemes /f, p, p^h, m, s, t, t^h, n/, three vowel phonemes /aː, i, u/, three codas /m, n/ and ∅, and six tones /55, 25, 33, 21, 23, 22/. This method produced 432 syllables (162 attested, 270 nonwords). 10 native speakers of Cantonese were asked to judge whether each stimulus was a word in Cantonese, as well as its wordlikeness rating on a 7-point scale. Conditional phonotactic probability (PP) and neighborhood density (ND) values were calculated for all items.

Results. Wordlikeness ratings did not always favor accidental over systematic gaps, as illustrated by Fig. 1. The difference between wordlikeness ratings for words and nonwords was significant ($U = 39217.5, p < 0.01$), as were the differences between labial/coronal ($U = 480.5$, $p = 0.01$) and coronal/onset-tone ($U = 238$, $p < 0.01$) gaps, but not between labial/onset-tone gaps ($U = 1405$, $p = 0.022$; $\alpha = 0.016$ for all tests). Multiple linear regression using ND and PP as factors captured a moderate amount of the overall variance ($R_a^2 = 0.3429$, $F(2, 429) = 113.4$, $p < 0.01$), with both factors emerging as significant. Where present, the trends relating phonotactic probability to goodness rating are weak, but lexical density bore a strong linear relation to wordlikeness for systematic gaps, particularly onset-tone gaps, as can be seen in Fig. 2.

Conclusions. Wordlikeness judgments were found to vary with lexical density and, to a lesser extent, conditional bigram probability. This is perhaps unexpected in light of work claiming a large role for PP in English [3] [7]. We argue this to be an effect of the relative phonotactic densities of Cantonese (Fig. 3) and English (Fig. 4): due to its distributon, PP is a much better cue to lexicality in English than in Cantonese. While regression analyses show a significant amount of the variability in judgments to be a function of ND and PP, judgment of systematic gaps also appears to be governed by categorical constraints. When compared to similar studies for languages such as English, these results suggest language-specific roles for phonotactic probability and neighborhood density with respect to their influence on the wordlikeness judgments of nonwords.

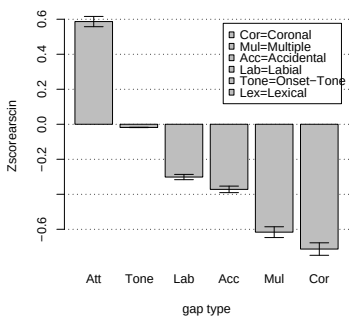

Figure 1: Arcsine-transformed goodness ratings by gap type. Error bars show 95% confidence interval for the mean.

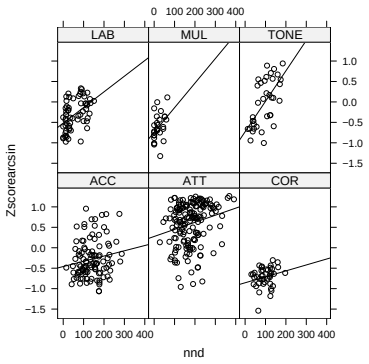

Figure 2: Wordlikeness as a function of lexical neighborhood density by syllable type.

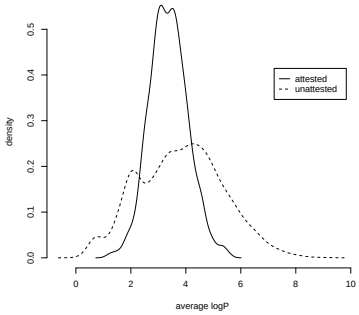

Figure 3: Kernel density plots of attested and unattested Cantonese monosyllables.

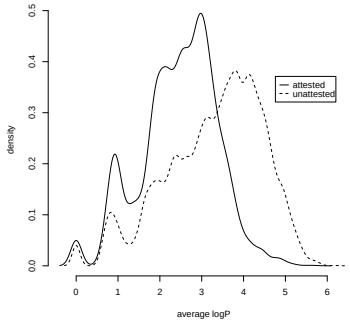

Figure 4: Kernel density plots of attested and unattested English monosyllables.

References. [1] T. Bailey and U. Hahn. Determinants of wordlikeness: phonotactics or lexical neighborhoods? *J. Memory & Language*, 44:568-591, 2001. [2] J. Coleman and J. Pierrehumbert. Stochastic phonological grammars and acceptability. In *Computation Phonology: ACL SIGPHON 3*, 49-56, Somerset, NJ, 1997. Association of Computational Linguistics. [3] S. Frisch, N. Large, and D. Pisoni. Perception of wordlikeness: effects of segment probability and length on the processing of nonwords. *J. Memory & Language*, 42:481-496, 2000. [4] M. Halle. Phonology in generative grammar. *Word*, 18:54-72, 1962. [5] E. Moreton. 2002. Structural constraints in the perception of English stop-sonorant clusters. *Cognition* 84:055-71. [6] J. Ohala and M. Ohala. Testing hypotheses regarding the psychological manifestation of morpheme structure constraints. In J. Ohala and J. Jager, eds., *Experimental Phonology*, 239-252. Academic Press, Florida, 1986. [7] M. Vitevitch, P. Luce, and D. Kemmerer. Phonotactics and syllable stress: implications for the processing of spoken nonsense words. *Language and Speech*, 40(1):47-62, 1997.