Perceptual Dispersion in the Lexicon
Peter Graff, MIT Department of Linguistics and Philosophy, graff@mit.edu, oral presentation

Introduction: We explore the hypothesis that the global organization of the lexicon maximizes the perceptual distinctness of words, by preferentially relying on highly perceptible contrasts, even when the phonotactics of the language permit less perceptible ones. We predict that distinctions among words will predominantly rely either on few highly perceptible contrasts (e.g. /tin/ vs. /km/ with a place contrast in prevocalic position) or on many globally distributed contrasts, with multiple differences keeping the words apart (e.g. /stis/ vs. /ʃaks/ with a postvocalic place contrast further disambiguated by other differences). We present evidence for this hypothesis from a study investigating the typology of minimal pairs cross-linguistically and a study relating the frequency of English minimal pairs to patterns of perceptual confusion.

Methods: Dictionaries of 20 languages from 15 families have been analyzed. To approximate the importance of a contrast in distinguishing words, we utilize the number of minimal pairs relying exclusively on that contrast (e.g. the pair rib/rip relies on word-final voicing) as our dependent measure in all mixed log-linear models (Poisson regression) reported below. To control for the effect of phonotactics, we include log-context-frequency (contrasts in frequent contexts distinguish more words), the entropy of the distribution of feature values (uniformly distributed features distinguish more words) as well as the maximal random effects structure justified by the data (intercepts and slopes grouped by language) in all models.

1. Place contrast in stops: In the 15 languages surveyed which, like English, have place contrasts for stops in intervocalic (append/attend), pre-vocalic (spy/sty), post-vocalic (sweeps/sweets) and inter-consonantal (lisps/lists) position, the number of words disambiguated solely by these contrasts decreases as a function of the perceptibility of cues to place (Fujimura et al. 1986, Ohala 1990), which decrease in the order of V_V>C_V>V_C>C_C ($\beta_{AVBCD}$=-0.43, $z=-2.58$, $p<.01$; $\beta_{ABvCD}$=-0.06, $z=-0.3$, $p=.76$; $\beta_{ABCvD}$=-1.17, $z=-5.11$, $p<.0001$, $\chi^2(3)= 25.66$, $p<.0001$; forward difference coding; Figure 1). This effect holds while controlling for frequency distributions of sounds and features, and language specific differences therein.

2. Phoneme confusability in English: We find that the number of minimal pairs based on consonantal differences in English increases significantly as the consonants are more perceptually distinct (a property inferred from data in Miller and Nicely 1954), while controlling for sound specific distributions ($\beta=0.04$, $z=2.81$, $p<.005$; $\chi^2(1)= 8.01$, $p<.005$; Figure 2).

Conclusion: These results evidence a link between distinctness of words in the lexicon and the perception of speech as indicated by the significant effect of perceptibility beyond the phonotactic control variables. They suggest that the lexicon preferentially assigns minimal pairs for any feature to contexts where that feature is better perceived. We discuss implications of these findings for competing theories of the relation between perceptual distinctiveness and phonological patterns (e.g. Ohala 1981, Steriade 2001, Blevins 2004, Hayes and Steriade 2004).
Figure 1: Number of minimal pairs for place contrasts in stops in inter-vocalic (V_V), pre-vocalic and post-consonantal (C_V), post-vocalic and pre-consonantal (V_C), and inter-consonantal context (C_C; from left to right) in languages that have place contrasts in all four contexts.

Figure 2: Minimal pairs for X and Y residualized against frequency and the entropy of their distribution (y-axis) increases with the perceptual distinctness of X and Y (Miller and Nicely 1954; x-axis).