Statistical testing of the functional load hypothesis and implications for phonological theory

**Keywords:** functional load, merger, language change

The idea that phonological neutralization is constrained by the communicative function of lexical contrast has a long history (e.g., Trubetzkoy 1939, Martinet 1955, Hockett 1967). A few case-studies have asked whether occurring neutralizations have a lower impact on lexical contrast than non-occurring ones, either by creating fewer homophones (e.g., Silverman 2010, Kaplan in press) or by incurring a relatively small drop in lexical-level entropy (Surendran and Niyogi 2006). Results have been mixed: some case-studies are consistent with the functional load account (Silverman 2010, Kaplan in press) while others contradict it (e.g., King 1967) or are judged inconclusive (e.g. Surendran & Niyogi 2006). However, if functional load is just one of many influences on sound change we would expect to find many individual ‘counter-examples’ to the functional load hypothesis. Consequently, we need a set of case-studies that can be assessed jointly using statistical methods. As part of an ongoing investigation of relationships between functional load and contrast neutralization, we are building a database of diachronically recent mergers in languages for which there exist frequency-coded phonemic word lists. The database currently includes data related to forty distinct mergers occurring in German, French, Korean, and several varieties of English. This data reveals a statistically significant difference between the distributions of these mergers versus comparable non-occurring mergers, in three measures of the functional load for phonemic oppositions: changes in homophony, lexical entropy and segmental entropy.

To ask whether a neutralized contrast had a relatively low functional load, comparison to some set of similar, but non-merged contrasts is necessary (Surendran & Niyogi 2006). Rather than attempting to draw a defensible distinction between plausible and implausible mergers in the construction of comparison sets, we included all structurally similar oppositions. For vowel-vowel mergers, for example, we used all pair-wise vowel-vowel mergers given the vowel inventory. The inclusion of phonologically implausible mergers is supported by Mann-Whitney/Wilcoxon Rank Sum tests, which show that for these languages, mergers between phonologically similar contrasts do not score lower on our measures of functional load than more phonologically distant contrasts. Figure 1 shows the ranks of each opposition in the data set with regard to the degree of associated homophony increase; occurring mergers are significantly more likely to be ranked lower than non-occurring mergers (p < .001). Ranking oppositions either by their associated decreases in lexical or segmental entropy upon merger (Surendran and Niyogi 2006) show similar statistically significant shifts. Thus, it could be the number of minimal pairs or simply the frequency of the contrasting phonemes (or both) that is responsible for the effect; our analysis cannot yet distinguish between the two possibilities.

The relationship between functional load and phonological change suggested by these results cannot be straightforwardly modeled in traditional generative phonology: current theories generally do not allow phonology to make reference to the lexicon, to minimal pairs or to phoneme frequency. In classical Optimality Theory, for example, the lexicon and the phonological grammar are completely separate: the principle of Richness of the Base requires that the locus of explanation for all phonological patterns be in the constraint ranking, not in the set of words the language happens to have. Dispersion Theory, which allows contrast maintenance to play a role in phonological grammars, explicitly rejects the possibility of considering contrasts among individual lexical items (Padgett 2003). Although this practice correctly rules out certain unattested patterns, it leaves the theory without a mechanism for explaining the relationship between the properties of the lexicon and phonological neutralization found here. We will argue that models in which phonological information is active at lexical and sublexical category levels (e.g., Bybee 2001, Pierrehumbert 2002, Blevins 2004, Wedel 2007) may provide better prospects for a satisfactory account of these data.
These results also highlight the importance of computational methods for phonological research: the computational tools used here reveal systematic phonological behavior that could not have been observed by traditional phonological analysis. Computational methods complement traditional ones and add to the phonologist's toolbox.

Figure 1

Legend: All phoneme oppositions in the database are ranked by proportional change in homophony upon merger. Red vertical lines indicate actual mergers. Median ranks for actual versus non-occurring mergers are 509 and 845 respectively. Mann-Whitney U = 44015; n1 = 40, n2 = 1676; p < .001, one tailed.

References