Coalescing sources of bias in perception:
Lexical and prelexical influences on the processing of phonological features

Alexander Martin & Sharon Peperkamp

Laboratoire de Sciences Cognitives et Psycholinguistique (ENS, EHESS, CNRS)
École Normale Supérieure – PSL Research University

Previous research in perception has demonstrated asymmetries in the processing of phonological feature contrasts during word recognition (Cole, Jakimik, & Cooper, 1978; Ernestus & Mak, 2004; Martin & Peperkamp, 2015). In French for example, Martin & Peperkamp (2015) found that changes in voicing are more likely to lead to a mispronunciation being recognized as a real word (e.g., /ɡolonɛl/ → /kolonɛl/) than changes in manner or place of articulation (/folonɛl/ or /tolonɛl/ → /kolonɛl/). The present study considers two possible sources of these asymmetries: top-down knowledge of the lexicon, and bottom-up phonetic perception.

First, we explore lexical sources of bias by measuring the functional load (henceforth FL) of the different phonological features. FL can be coarsely defined as the weight of a certain contrast in a given phonological system. As our question requires the quantification of the differences between the individual features, the common practices of measuring information theoretic entropy or the number of minimal pairs—both of which yield a single number for each feature (e.g., Surendran & Niyogi, 2003, 2006; Wedel, Kaplan, & Jackson, 2013)—are deemed insufficient. We propose a new measure of FL based on minimal pair (henceforth MP) counts, which, in addition to stripping away effects derived from phonotactics and individual phoneme frequency, yields a distribution of scores for each feature, allowing the use of inferential statistics. We calculate a score $r_{ij}$ for each phoneme ($i$)/feature ($j$) combination by iterating over the lexicon. The method consists in calculating the number of observed MPs ($o$) over the number of possible MPs ($p$): $\log (o_{ij}/p_{ij})$. For example, when we are looking at the place feature and the phoneme /p/, thus $r_{p/\text{place}}$ upon encountering the word /po/ (“skin”), we know that an MP is theoretically possible (i.e., a change in place on the segment /p/ yields the possible word /to/). Thus $p = 1$. Furthermore, we know that the word /to/ does exist (taux, “amount”). Thus $o = 1$. If we next consider a case such as /propr/ (propre, “clean”), we observe that the theoretical MP it would form with a place change is possible (i.e., /tɔpr/ is a phonotactically legal word). Thus $p = 2$ now (the scores are cumulative as we iterate over the lexicon). However, this word does not actually exist in French, and $o$ therefore remains at 1. Now if we consider the case of /pltɛ/ (pluie, “rain”), we know that the theoretical MP it would form from a place change is not possible (i.e., /tlɛ/ does not follow French phonotactics), thus $p$ remains at 2. Of course if an MP is not possible, it will not be observed, and indeed $o$ remains at 1. This process is repeated over the entire lexicon for each phoneme/feature combination to obtain distributions; see Fig. 1. When applied to the nouns of the French lexicon—the class of words used in Martin & Peperkamp (2015)—our measure reveals that place differences yield significantly more minimal pairs than manner or voicing differences ($F = 9.1, p < 0.001$).

Turning to perception, it is important to have a baseline measure of how different individual features are perceived outside of lexical context. Listeners, however, are very good at perceiving the sounds of their native language, and in identification tasks regularly achieve ceiling performance (e.g., Hall & Hume, 2013; Plauché, Delogu, & Ohala, 1997). To get around this, some recent work has used explicit similarity judgments to assess the question of similarity (Bailey & Hahn, 2005; Hahn & Bailey, 2005). These studies revealed that the number of featural differences between two sounds is predictive of their perceived similarity, but did not address the question of the (dis)similarity of the features amongst themselves.

Here, we designed a discrimination experiment that does not give rise to ceiling performance. In an ABX task, participants were presented trials of the form A: /papaba/ - B: /pababa/ - X: /papaba/ and asked whether the X stimulus was identical to A or B. The stimuli were generated using three different voices of the Apple Say speech synthesizer. We manipulated the position of the difference (first, second, or third), the surrounding vowels (which were always identical, /a/, /i/ or /u/), and crucially, the
phonological feature. Participants were therefore presented voicing trials, but also manner trials such as /tisi/ /tisi/ /tisi/ (correct response B) and place trials like /tisi/ /tisi/ /tisi/ (correct response A). We measured both accuracy and response times, and analyzed all results using mixed-effects models. The results with 48 French participants revealed that for manner contrasts, participants were both more accurate ($\beta = -0.82, p < 0.0001$), and, on correct trials, faster to give their responses ($\beta = 0.07, p = 0.025$) than for place or voicing contrasts; see Fig. 2.

Overall, it is clear that asymmetries can be observed amongst otherwise fully contrastive features on multiple levels. Recalling Martin & Peperkamp (2015)’s results for French obstruents, as shown in (1c), we propose that the differences they observed in lexical perception (i.e., that voicing was less important than manner or place) can be explained by coalescing information from two levels. First, all listeners are subject to bottom-up acoustic biases during speech perception. Our prelexical perception experiment with French listeners revealed that manner differences were perceived better than differences in place or voicing; see (1a). This may be based in the very stark acoustic differences that distinguish stops from fricatives, and can explain why listeners in Martin & Peperkamp (2015)’s study had increased difficulty in recognizing manner mispronunciations as real words. Second, listeners have specific knowledge of their lexicon, and may lend more attention to cues related to contrasts that distinguish many words in their native language. Our measure of this functional load in French revealed that place contrasts are significantly more common than manner or voicing contrasts; see (ib). If listeners are biased by this knowledge, then it would explain why participants in Martin & Peperkamp (2015)’s experiment had increased difficulty in recognizing place mispronunciations as real words. In this view, the levels described in (1a) and (1b) add up to what we observe in (1c), with the lexicon and low-level perception biasing listeners towards paying more attention to place and manner.

Clearly, perception is modulated by various sources, and our results, though they focus on a subset of the French phonological inventory, demonstrate how the processing of phonological features in one’s native language is influenced both by low-level perception and by lexical knowledge.

(1) a. Prelexical perception: MANNER > PLACE, VOICING
b. Lexicon: PLACE > MANNER, VOICING
c. Word recognition: PLACE, MANNER > VOICING

Figure 1: Functional load of phonological features in French nouns (in log space). Values closer to zero represent higher FL.

Figure 2: Mean error rates (white bars) and response times (black bars) in our prelexical ABX task. Error bars represent SEM by participant.

References


