Journal of Phonetics 62 (2017) 1–11

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/Phonetics

Research Article

Assessing the distinctiveness of phonological features in word recognition: Prelexical and lexical influences

Alexander Martin *, Sharon Peperkamp

Laboratoire de Sciences Cognitives et Psycholinguistique (ENS, EHESS, CNRS), Département d'Études Cognitives, École Normale Supérieure – PSL Research University, 29 rue d'Ulm, 75005 Paris, France

ARTICLE INFO

Article history:
Received 30 May 2016
Received in revised form 17 January 2017
Accepted 24 January 2017
Available online 16 February 2017

Keywords:
Word recognition
Phonological features
Perceptual similarity
Lexicon
Functional load

ABSTRACT

Phonological features have been shown to differ from one another in their perceptual weight during word recognition. Here, we examine two possible sources of these asymmetries: bottom-up acoustic perception (some featural contrasts are acoustically more different than others), and top-down lexical knowledge (some contrasts are used more to distinguish words in the lexicon). We focus on French nouns, in which voicing mispronunciations are perceived as closer to canonical pronunciations than both place and manner mispronunciations, indicating that voicing is less important than place and manner for distinguishing words from one another. We find that this result can be accounted for by coalescing the two sources of bias. First, using a prelexical discrimination paradigm, we show that manner contrasts have the highest baseline perceptual salience, while there is no difference between place and voicing. Second, using a novel method to compute the functional load of phonological features, we show that the place feature is most often recruited to distinguish nouns in the French lexicon, while voicing and manner are exploited equally often.

1. Introduction

What makes two words sound similar to each other? Consider the English word *pin* – /pɪn/. Intuitively, we can understand how a word like *shin* – /ʃɪn/ sounds more similar to *pin* than a word like *train* – /tɹe͡ɪn/ does. Indeed *pin* and *shin* form a minimal pair; the two words are minimally different, in that they share all but one phoneme. Yet cross-modal priming experiments have shown that a word like *bin* – /bɪn/, which also forms a minimal pair with *pin*, more strongly activates *pin* than *shin* does (e.g., Connine, Blasko, & Titone, 1993; Milberg, Blumstein, & Dworetzky, 1988). This is because the segments that distinguish *pin* from *shin* share fewer phonological features than those that distinguish *pin* from *bin*. Now consider the word *tin* – /tɪn/. Both the /t/ in *tin* and the /b/ in *bin* are one feature different from the /p/ in *pin* (a difference in place and voicing[1] respectively). Is the nature of the featural difference pertinent for the notion of similarity?

Research on lexical perception has demonstrated that featural differences in one's native language are not all perceived as equally distinct. In both English (Cole, Jakimik, & Cooper, 1978) and Dutch (Ernestus & Mak, 2004), mispronunciations have been shown to be less disruptive for word recognition (i.e., easier to recognize) if they involve a change in voicing than if they involve a change in place or in manner. This indicates that a difference in voicing is perceived as less stark than a difference in another major class feature in these languages. More recently, Martin and Peperkamp (2015) exposed French listeners to a series of auditorily- (or audiovisually-) presented nouns supposedly produced by a stroke patient. These included correctly pronounced words, mispronounced words, and non-words that did not resemble any real word. Participants were asked to press a button when they recognized a word – whether it was correctly pronounced or mispronounced – and report it. All mispronunciations involved a change in one of the major class features: voicing, manner, or place on a word-initial obstruent. The results from the audio-only version of that experiment, reported as the proportion of correctly identified mispronounced words, are reproduced in Fig. 1.[2] Similar

* Corresponding author.
E-mail address: alxndr.martin@gmail.com (A. Martin).

[1] Note that throughout this paper we will refer to any two-way laryngeal contrast as "voicing", although the phonetic realization of this contrast may vary across languages.

[2] The results were not significantly different by modality.

http://dx.doi.org/10.1016/j.wocn.2017.01.007
0095-4470/

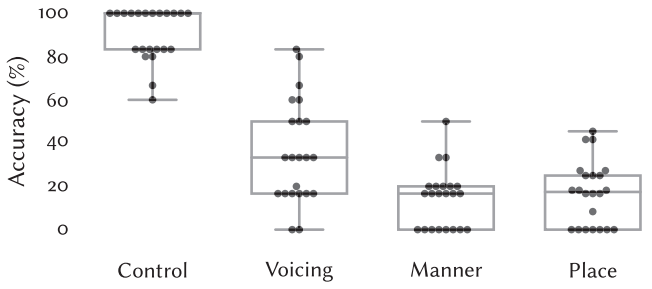

Fig. 1. Boxplot of participant means from the audio-only version of the mispronunciation detection task reported in Martin and Peperkamp (2015) by condition. The central line in the boxplot represents the median; the space between the central line and the bottom or top of the box represents the second and third quartile spread; and the distance from the bottom or top of the box to the tip of the whiskers represents the first and fourth quartile spread. In the dotplots, each dot represents an individual's score.

to the previous findings for English and Dutch (Cole et al., 1978; Ernestus & Mak, 2004), words with a voicing mispronunciation were more likely to be recognized than those with a manner or a place mispronunciation. For example, the word *sommet*, /sɔmɛ/ – "summit" was more likely to be recognized when it was presented as /zɔmɛ/, with a mispronunciation in voicing, than when it was presented as /fɔmɛ/ or /tɔmɛ/ (a place or manner mispronunciation, respectively). Thus, the voicing feature's role in contrasting words from one another is perceived as different than that of the other features.

The sources of this asymmetry remain unclear, however, and could be due to a number of factors. Most obvious is the acoustic proximity of the sounds being considered. Some sounds are acoustically closer, and will thus be perceived as more similar than other, distant, sounds. A further source of bias is language-specific knowledge. Listeners may use knowledge of their native language for the purposes of efficient word recognition. That is, they may preferentially attend to cues associated with featural contrasts which are more informative in their language. Indeed, listeners are influenced both by acoustic information and by language-specific knowledge (Ernestus & Mak, 2004; Johnson & Babel, 2010). Ernestus and Mak (2004), for example, argued that Dutch listeners are influenced by a process of initial fricative devoicing in their language, which renders voicing information on these segments uninformative. This would explain why these listeners ignore voicing mispronunciations more often than manner mispronunciations in a lexical decision task. Similarly, Johnson and Babel (2010) found language-specific influence using a similarity judgment task. They had native English- and Dutch-speaking participants rate the similarity of pairs of VCV non-words containing English fricatives, and showed that Dutch listeners rated [s], [ʃ], and [θ] as more similar to each other than English listeners did. They argued that this is due to the phonological status of these sounds in the respective languages. While all three sounds are distinctive in English, [ʃ] and [θ] are not phonologically distinctive in Dutch; the former is a contextual allophone of /s/ and the latter does not occur at all. However, in an AX discrimination experiment, Dutch listeners' response times were not shown to differ from English listeners'; both groups discriminated the same pairs of sounds equally well. The authors argued that their discrimination task reveals low-level acoustic differences between the stimuli, with some of the contrasts yielding longer response times because of their acoustic proximity (e.g., [f]∼[θ] and [h]∼[x]), regardless of the native language of the listener, while their similarity judgment task reveals language-specific influences, with Dutch listeners being perturbed by the absence of [θ] and [ʃ] as phonemes in their native language.

Note, though, that this reasoning does not explain why in English and French, voicing mispronunciations are also harder to detect (Cole et al., 1978; Martin & Peperkamp, 2015), because the voicing feature is fully distinctive in these languages (voicing contrasts can be neutralized in English and French but never word-initially). These results do not necessarily imply that listeners are not influenced by lexical patterns during word recognition. Indeed, following, inter alia, Hall (2013), we argue that a more gradient understanding of "distinctiveness" is necessary to properly address this issue. If, for example, there were *fewer* voicing minimal pairs than place and manner minimal pairs in English and French, this could explain why words presented with voicing mispronunciations were perceived as closer to the target word. Here, we further explore gradient distinctiveness using a combination of experimental and computational techniques.

The specific aim of our research is to disentangle low-level, prelexical influences from top-down, lexical ones in word recognition. To this end, we take French obstruents as a case study, allowing for a direct comparison with the results on lexical perception from the mispronunciation detection task reported in Martin and Peperkamp (2015), which we take as our starting point. Building on those results, we start off with an examination of the way phonetic differences between features are perceived outside of lexical context, using a prelexical discrimination task. We then examine the French lexicon by measuring the *functional load* of various feature contrasts as a proxy for the lexical knowledge shared by speakers of French. This allows us to understand if there are asymmetries in the usage of these different features, even though they are not affected by any phonological process. Finally, we compare our results with the word recognition results reported in Martin and Peperkamp (2015), and propose that the relative weight of phonological features during word recognition is determined jointly by the role of these features in both bottom-up acoustic perception and top-down lexical knowledge.

2. Prelexical perception

The perceptual similarity of speech sounds has been investigated for decades, focusing mostly on the effects of different types of noise on perceptual confusion (e.g., Bell, Dirks, & Carterette, 1989; Cutler, Weber, Smits, & Cooper, 2004; Miller & Nicely, 1955; Weber & Smits, 2003). For instance, Miller and Nicely (1955) presented a series of English syllables embedded in different kinds of noise (including low-pass filtering and white noise) at various signal-to-noise ratios (SNR) and asked participants to report what consonant the syllable began with. They found that *place* of articulation was more likely to be confused than *voicing*, across consonants and across different SNRs. While this line of research is important for understanding speech perception in noisy conditions, it cannot provide us with an accurate *baseline* of perceptual similarity of speech sounds, because noise affects individual