Research Article

Non-native phonemic discrimination, phonological short term memory, and word learning

Noah H. Silbert [a,*], Benjamin K. Smith [b], Scott R. Jackson [b], Susan G. Campbell [b], Meredith M. Hughes [b], Medha Tare [b]

[a] University of Cincinnati, Department of Communication Sciences & Disorders, P.O. Box 670379, Cincinnati, OH 45267, USA
[b] University of Maryland, Center for Advanced Study of Language, 7005 52nd Avenue, College Park, MD 20742, USA

ARTICLE INFO

ABSTRACT

Accurately perceiving non-native speech sounds is known to be very difficult. Numerous studies provide strong and converging evidence that this difficulty varies systematically, depending on the properties of the non-native sounds and the native language of the listener. There is substantially less research on how phonetic and phonological structure relates to individual differences in the ability to perceive non-native phonemes, though individual differences in auditory abilities are well-documented. The present work reports two experiments aimed at elucidating the structure of individual differences in non-native speech perception and the relationship between these abilities, phonological short term memory, and early second language word learning. In Experiment 1, confirmatory factor analysis of discrimination data for nine non-native contrasts from different languages indicates that voicing and place (segmental) contrasts pattern together and distinct from tone (suprasegmental) contrasts. In Experiment 2, the results indicate that phonological short term memory and discrimination ability both predict word learning accuracy and that discrimination ability does so in a mostly feature-specific manner.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Non-native speech perception is beset by a number of difficulties, among them the fact that the speech sounds in a non-native language may bear complicated relationships to speech sounds in a listener's native language. For example, two phonologically distinct sounds in a second language may sound, to varying degrees, like better or worse variants of a single native category, as in the well-studied case of Japanese native speakers' difficulty in accurately perceiving English /r/ and /l/ (e.g., Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2004; Flege, Takagi, & Mann, 1996; Lively, Pisoni, Yamada, Tohkura, & Yamada, 1994).

A second-language learner could benefit substantially from the ability to accurately perceive the acoustics of non-native speech sounds regardless of their correspondence to native categories. This ability is likely to be particularly beneficial at early stages of second language acquisition, before the learner has acquired knowledge of higher-level linguistic structure that could help disambiguate words that would otherwise be functionally homophonous to the non-native listener. Conversely, if a learner cannot reliably discriminate a pair of sounds that are phonologically distinctive in a second language, words that rely on this difference to cue distinct meanings will be difficult, if not impossible, to perceive and store in memory accurately.

Prior research has shown that some non-native speech sounds are more difficult than others to perceive accurately, primarily due to differences in the mapping between native and non-native phonological categories (e.g., Best, Hallé, Bohn, & Faber, 2003; Best, McRoberts, & Goodell, 2001; Flege, 1993, 1995; Flege, Schirru, & MacKay, 2003). However, relatively few studies have focused on individual differences in the ability to perceive non-native phones accurately, though individual differences in auditory abilities are well-documented (Humes, Kidd, & Lentz, 2013; Kidd, Watson, & Gygi, 2007; Surprenant & Watson, 2001). Recent work has shown that individuals differ in perceptual training outcomes (Golestani & Zatorre, 2009) and that individual differences in certain auditory abilities can play an important role in determining the optimal structure of perceptual training protocols (Perrachione, Lee, Ha, & Wong, 2011). Work on individual differences in second language speech perception has often been carried out in the context of

measures of aptitude for early second language learning (e.g., Carroll, 1962, 1990; Pimsleur, Sundland, & McIntyre, 1964). Because of the relative paucity of research on individual differences in phonemic discrimination ability, very little is known about the structure of this ability.

We describe two experiments designed to investigate individual differences in non-native perceptual abilities. The first experiment was designed to elucidate the structure of non-native phonemic discrimination ability by comparing different factor analysis models of these abilities fit to data from non-native discrimination of contrasts from nine languages. The results of this experiment indicate that non-native phonemic discrimination abilities are multidimensional, reflecting segmental and suprasegmental phonological structure. The second experiment was designed to establish the degree to which phonemic discrimination ability might influence early word learning (as measured by a paired associates task) while controlling for phonological short term memory. The results of the second experiment indicate that discrimination ability predicts word learning accuracy in a largely feature-specific manner, even when taking phonological recognition memory into account.

1.1. Speech learning, perceptual assimilation, and native language magnets

Within the first year of life, infants begin to lose the ability to discriminate speech sounds that are not distinctive in their speech environment (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Werker, Gilbert, Humphrey, & Tees, 1981; Werker & Tees, 1984). As phonological categories develop, listeners have more difficulty discriminating sounds that, while acoustically distinct, fall within a single category (Iverson & Kuhl, 1995, 2000; Kuhl, 1991). Recent work has probed the neural underpinnings of phonological category development and this native-language "magnet effect," the idea that perceptual space is warped near the center of native-language phonetic categories such that the similarity of acoustically distinct stimuli varies with the distance from the category center (e.g., Kuhl et al., 2008).

When adults learn a second language, fully developed native categories can influence the learning of non-native categories in complicated ways. Two conceptually related models have been developed to document and explicate these interactions, each with a slightly different focus. The Speech Learning Model (SLM) of Flege and colleagues (Flege & Eefting, 1987b; Flege et al., 2003; Flege & Wang, 1990) focuses on the relationships between native and non-native categories in speech production and perception and the effects of age of acquisition, extensive exposure, and formal schooling on these relationships. The Perceptual Assimilation Model (PAM) of Best et al. (2001) provides a typology of mappings between native and non-native categories, relating assimilation of non-native sounds to native categories and relative discriminability of pairs of non-native sounds.

Work on the SLM provides extensive evidence of perceptual differences between native and non-native listeners and between less and more experienced non-native listeners for a variety of phonological categories and for a number of combinations of native and non-native languages. Numerous studies show that native speakers of English, on the one hand, and native speakers of Spanish, Mandarin, Arabic, French, and Dutch, on the other, perceive voicing distinctions (e.g., /t/ vs /d/) differently in a variety of contexts (Flege & Bohn, 1993; Flege & Eefting, 1987a, 1987b; Flege, Munro, & Skelton, 1992; Flege & Schmidt, 1995; Flege & Wang, 1990; Flege, 1984, 1989, 1993; MacKay, Flege, Piske, & Schirru, 2001), and that Japanese native listeners and English native listeners perceive English consonants, including, but not limited to /r/ and /l/, differently (Flege et al., 1996; Guion, Flege, Akahane-Yamada, & Pruitt, 2000). Similarly, a number of studies show differences in the perception of English vowels by native speakers of English, on the one hand, and native speakers of German, Spanish, Mandarin, Korean, Italian, and Japanese, on the other (Bohn & Flege, 1990; Flege, Bohn, & Jang, 1997; Flege, MacKay, & Meador, 1999; Flege, Munro, & Fox, 1994; Flege, 1991; Fox, Flege, & Munro, 1995; Freida, Walley, Flege, & Sloane, 1999; Hojen & Flege, 2006; Nozawa & Flege, 2000). There is ample evidence that English native speakers also differ from native speakers of various other languages with respect to the perception of phones from those languages (e.g., Cheon & Anderson, 2008; Francis & Nusbaum, 2002; Hao, 2012; So & Best, 2010; Sun & Huang, 2012; Wayland & Li, 2008). The general pattern observed across languages and phonological categories is that, although non-native perception often improves with experience, even listeners with extensive exposure to and/or intensive study of a second language differ from native speakers, typically exhibiting lower accuracy and distinctly non-native patterns of perception and categorization of various phones.

Work in the PAM framework provides strong evidence that the manner in which non-native sounds map onto native sounds influences discriminability of the non-native sounds. When two non-native phones are assimilated similarly to a single native category, discrimination is less accurate than when they are assimilated asymmetrically. Discrimination in the asymmetric, single-category assimilation case is in turn less accurate than when each of two non-native phones is assimilated to a separate native category. In some cases, non-native phones are not assimilated to native categories at all, in which case listeners may be able to discriminate non-native phones fairly well by attending directly to acoustic differences. This general pattern holds for consonants (e.g., Best et al., 2001; Best, McRoberts, & Sithole, 1988; Polka, 1992), vowels (e.g., Best et al., 2003), and tones (e.g., So & Best, 2010).

Both the SLM and the PAM frameworks are focused on how phonetic and phonological factors influence the relative difficulty of perceiving (and, in the case of SLM, producing) non-native contrasts. Work in these frameworks has revealed a great deal about how speech behavior varies with experience, and how the mapping between native and non-native phones influences discriminability, but substantially less is known about how the ability to discriminate non-native phones varies across individual listeners or how such abilities are related to learning of higher-level linguistic units (e.g., words) or global constructs like language learning aptitude.