ICA 2013 Montreal
Montreal, Canada
2 - 7 June 2013

Speech Communication
Session 5aSCb: Production and Perception II: The Speech Segment (Poster Session)

5aSCb23. Phonological structure, non-native phoneme discrimination, working memory, and word learning
Noah Silbert*, Benjamin K. Smith and Scott R. Jackson

*Corresponding author's address: Center for Advanced Study of Language, University of Maryland, College Park, MD 20742, nsilbert@umd.edu

It is well known that perception of non-native speech sounds is influenced by exposure and the mapping between non-native and native phonological categories. However, very little is known about the relationships between phonological structure, individual differences in non-native phoneme discrimination ability, and non-native word learning. These relationships are important in the design of tests for personnel selection for second language training. Two experiments were conducted to probe the generality of phoneme discrimination ability and the role of phonological structure and discrimination ability in word learning. In one experiment, 169 participants discriminated non-native contrasts from nine languages - three voicing/laryngeal contrasts, three place contrasts, and three tone/intonation contrasts. Confirmatory factor analysis model comparisons show that correlations between discrimination accuracies across contrasts are driven by low-level phonological structure (featural and segmental/super-segmental properties). In a second experiment, phonological working memory and voicing, place, and tone discrimination were measured for 167 participants and used to predict learning of pairs of non-native words differing in voicing, place, and tone. Consistent with the results from the first experiment, discrimination ability predicts accuracy in word learning above and beyond the ability of phonological working memory and according to feature-specific differences.

Published by the Acoustical Society of America through the American Institute of Physics
Second Language Speech Perception

Native language speech perception is complex and only partially understood. Second language speech perception is beset by a number of additional difficulties, chief among them the fact that the distinctive speech sounds in a second language often bear complicated relationships to the distinctive speech sounds of a listener’s native language. A substantial body of research documents the relationships between non-native speech perception, non-native speech production, and the relationships between native and non-native phonetic categories (e.g., Best, Hallé, Bohn, & Faber, 2003; Best, McRoberts, & Goodell, 2001; Flege, Schirru, & MacKay, 2003; Flege, 1993, 1995; Guion, Flege, Akahane-Yamada, & Pruitt, 2000; Hao, 2012; MacKay, Flege, Piske, & Schirru, 2001; So & Best, 2010). However, very little work has been done relating low-level second language perceptual abilities and global measures of aptitude or proficiency (e.g., Carroll, 1962, 1990; Pimsleur, Sundland, & McIntyre, 1964), and much of what little work has been done in this area provides mixed evidence for the predictive utility of discrimination ability in tests of aptitude or proficiency. There is evidence that poor phonemic discrimination ability reliably predicts underachievement in early foreign language learning (Pimsleur et al., 1964). There is also evidence that discrimination performance is strongly correlated with short-term memory and phonetic coding, which is the association of phones with arbitrary symbols (Harnsberger, 2000), both of which are good predictors of proficiency (Bunting et al., 2011; Carroll, 1962, 1990).

Measuring Non-Native Phoneme Discrimination Ability

We describe two experiments that support the development and preliminary validation of a new measure of phonemic discrimination ability. In Experiment 1, we probe English native listeners’ discrimination of contrasts from nine languages in order to determine if individuals differ with respect to a single phonemic discrimination ability or multiple such abilities and, if the latter, how these abilities are related to each other. In Experiment 2, we seek to establish the degree to which phonemic discrimination ability predicts word learning while controlling for general phonological working memory ability.

In both experiments, we employed a series of tests of discrimination, each using pairs of sounds from a different language. The test employs the so-called ‘oddball’ or ‘triangle’ task (Ennis & Mullen, 1986; Flege, 2003). In each trial of an oddball task, the participant hears three words, two of which are distinct tokens of a word representing one phonological category (e.g., ‘bit’), and one of which is a token of a word that only differs from the other two with respect to a single contrast of interest (e.g., ‘pit’). The participant identifies which word differs from the other two. Each oddball test probed discrimination of a phonological contrast from one of nine languages. The nine languages/contrasts probed are shown in Table 1, along with (a subset of the) acoustic cues to the contrast, the primary cue type (temporal vs. spectral), and the phonological level of the contrast (segmental vs. suprasegmental).

**TABLE 1.** Languages and contrasts. Each contrast was produced by a native speaker in four word pairs. For each distinction, the Contrast column provides a phonetic description, Cues describes acoustic-phonetic cues, Primary Cue Type indicates whether the primary cue is temporal or spectral, and Phonological Level indicates whether the distinction is segmental or suprasegmental.

<table>
<thead>
<tr>
<th>Language</th>
<th>Contrast</th>
<th>Cues</th>
<th>Primary Cue Type</th>
<th>Phonological Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hindi</td>
<td>pre-voiced vs. voiceless</td>
<td>VOT, voicing during closure</td>
<td>temporal</td>
<td>segmental</td>
</tr>
<tr>
<td>Mandarin</td>
<td>unaspirated vs. aspirated</td>
<td>VOT, noise characteristics</td>
<td>temporal</td>
<td>segmental</td>
</tr>
<tr>
<td>Korean</td>
<td>lenis vs. aspirated</td>
<td>VOT, F0 (fundamental frequency)</td>
<td>temporal</td>
<td>segmental</td>
</tr>
<tr>
<td>Polish</td>
<td>retroflex vs. palatal</td>
<td>spectral shape, formant transitions</td>
<td>spectral</td>
<td>segmental</td>
</tr>
<tr>
<td>Arabic</td>
<td>velar vs. uvular</td>
<td>spectral shape, formant transitions</td>
<td>spectral</td>
<td>segmental</td>
</tr>
<tr>
<td>Igbo</td>
<td>bilabial vs. labio-velar</td>
<td>spectral shape, amplitude</td>
<td>spectral</td>
<td>segmental</td>
</tr>
<tr>
<td>Yoruba</td>
<td>tone: mid vs. high/low</td>
<td>F0 level and contour</td>
<td>spectral</td>
<td>suprasegmental</td>
</tr>
<tr>
<td>Japanese</td>
<td>pitch accent: HL vs. LH²</td>
<td>F0 level and contour</td>
<td>spectral</td>
<td>suprasegmental</td>
</tr>
<tr>
<td>Thai</td>
<td>tone: mid vs. low</td>
<td>F0 level and contour</td>
<td>spectral</td>
<td>suprasegmental</td>
</tr>
</tbody>
</table>

1 VOT = voice onset time, a measure of the time between release of an oral constriction and the beginning of periodic vocal fold oscillation.
2 HL = High-Low; LH = Low-High. These are the contrastive pitch accent patterns in Japanese. ‘H’ indicates the location of a F0 peak in a word.
The Structure and Predictive Utility of Phoneme Discrimination Abilities

One experiment was aimed at establishing if there is a single, general non-native phoneme discrimination ability or if there are multiple non-native phoneme discrimination abilities structured along cue- or feature-based dimensions. In order to elucidate the nature of phonemic discrimination ability (or abilities), we administered the battery of nine discrimination tests described above. Data analysis is currently underway, with initial confirmatory factor analysis (CFA) model fits and comparisons indicating that feature- and segment-based models provide better accounts of the data than do spectro-temporal or general-ability-based models.

The second experiment was aimed at estimating the degree to which discrimination abilities predict foreign language word learning above and beyond the predictive utility of phonological working memory. This experiment consisted of three oddball blocks, a test of phonological working memory, and three word-learning blocks. Data analysis is currently underway, with initial multilevel generalized linear model fits indicating (a) that discrimination ability predicts word learning accuracy even when taking phonological working memory into account, and (b) that discrimination ability predicts word learning accuracy in a feature-specific manner (e.g., the ability to discriminate place contrasts predicts the ability to learn words that differ with respect to place but not voicing or tone).

Conclusion

Initial model fits and comparisons suggest that (a) there are substantial, measurable individual differences in non-native phoneme abilities, (b) the ability to discriminate non-native phonemes is multidimensional and structured along phonological dimensions, and (c) above and beyond the predictive contribution of phonological working memory, the ability to discriminate a given non-native contrast predicts accuracy in learning words that differ with respect to that contrast but not other non-native contrasts. Pending the results of full analyses, these initial findings suggest that tests of non-native phoneme discrimination ability should include contrasts from multiple languages and that such tests are likely to contribute to the assessment of second language aptitude and proficiency.

ACKNOWLEDGMENTS

This work benefited from the contributions of Susan G. Campbell, Meredith M. Hughes, and Medha Tare.

REFERENCES


