Learning to Use Intonational Cues in Second-Language Speech Segmentation

Annie Tremblay
University of Kansas

Intonational cues such as fundamental frequency (F0) signal word boundaries differently across languages: In French and Korean, an F0 rise tends to signal word-final boundaries (Jun, 1998; Welby, 2006), whereas in English and Dutch, an F0 rise will instead signal word-initial boundaries (Beckman, 1986; Gussenhoven, 2004). Consequently, listeners from different language backgrounds use F0 differently to segment speech into words, with French and Korean listeners using F0 rise as a cue to word-final boundaries (Kim, Broersma, & Cho, 2012; Kim & Cho, 2009; Tyler & Cutler, 2009) and English and Dutch listeners using F0 rise as a cue to word-initial boundaries (Tyler & Cutler, 2009). Learning to segment speech in a second language (L2) may thus entail learning new associations between F0 cues and word boundaries. Can L2 learners accomplish this task? One prediction might be that it is easier for L2 learners to associate F0 cues to word boundaries in the L2 if the prosodic structures of the first language (L1) and L2 are similar. For example, in French and Korean, prominence is phrasal rather than lexical, with the basic tonal pattern of the Accentual Phrase (AP) being LHLH(*) and thus with the phrase-final H(*) signaling word-final boundaries (Jun, 1998; Jun & Fougeron, 2002; Welby, 2006). In English and Dutch, however, prominence is both phrasal and lexical, with most nouns being stressed word-initially (Cutler & Carter, 1987; Vroomen & de Gelder, 1995) and being signaled with a word-initial F0 rise when pitch accented (Beckman, 1986; Gussenhoven, 2004). Hence, one might predict that it would be easier for Korean listeners than for English or Dutch listeners to learn to segment French, and that English and Dutch listeners would make similar use of F0 cues to word boundaries in French.

Crucially, these predictions are not borne out. First, English-speaking L2 learners of French matched in French proficiency and experience with Korean-speaking L2 learners of French can learn to use F0 rise to segment French speech; in sharp contrast, Korean-speaking L2 learners of French cannot do so (Tremblay, Broersma, Coughlin, & Choi, 2016). This inability stems from the fine-grained prosodic differences between Korean and French: In Korean, the F0 peaks and lowers within the AP-final syllable such that it is already low by the beginning of following AP (Jun, 1998), whereas in French, the F0 peaks at the end of the AP-final syllable and lowers at the beginning of the following AP (Welby, 2006). The F0 rise in French thus occurs slightly too late for Korean listeners to be able to use it. Importantly, the similarity between the French and Korean prosodic systems creates a learning problem for Korean listeners, who assimilate the French system to their own and, accordingly, have difficulty learning to use F0 cues to word-final boundaries in French (see also Best & Tyler, 2007; Flege, 1995). Second, English-speaking L2 learners of French make lesser use of F0 cues in French word recognition than Dutch-speaking L2 learners of French matched with them in French proficiency and experience. This English disadvantage is attributed to the weaker functional load of F0 cues in English than in Dutch, with lexical stress being more strongly correlated with segmental information (i.e., vowel reduction) in English than in Dutch, and thus with English listeners relying less on F0 cues to word boundaries in French than Dutch listeners (see also Cooper, Cutler, & Wales, 2002; Cutler, Cooper, Wales, & Janssen, 2007).

These findings suggest that the L1 modulates the ease with which listeners learn to use intonational cues such as F0 rise in L2 speech segmentation, but does so in a complex way. An explanatory theory of the learning of L2 intonational cues must take into consideration not only the prosodic structures of the L1 and the L2, but also the fine-grained prosodic differences between the L1 and L2 systems (including their perceptual saliency) and the functional load of prosodic cues in the L1 (and L2).