Danica Reid
Simon Fraser University

Acoustic cues used by learners of English

Second language learners (L2) must acquire the ability to use acoustic and phonotactic information to segment long streams of continuous speech into meaningful words. In L2 speech segmentation, learners have two available options: adapt a strategy from their first language (L1) into their L2, or learn a new strategy based on the L2 input. The goal of this study was to identify how a listener's L1 segmentation strategies influence their real-time processing of potentially ambiguous L2 phrases. Two types of *s*C clusters were used to test L2 English speakers on their ability to segment fluent English speech. Cross-boundary *s*C clusters, where the word boundary falls between the two elements of the cluster (i.e., *keeps canning*) and word-initial *s*C clusters, where the word boundary precedes the /s/ (i.e., *keep scanning*) are distinguished by two inversely related cues in English: /s/-duration and allophonic aspiration. In cross-boundary *s*C clusters /s/-duration is shorter, and allophonic aspiration is present whereas in word-initial clusters /s/-duration is longer and allophonic aspiration is absent. Previous research on L2 English speakers from language backgrounds without systematic aspiration, Spanish (Altenberg, 2005) and French (Shoemaker, 2014), showed that identification of cross-boundary clusters was near chance. L2 English speakers with a L1 like Japanese, which has allophonic aspiration, were better able to make use of aspiration in the identification of *s*C cluster types (Ito & Strange, 2009). These studies showed that the segmentation strategies available from a L1 clearly have an impact on L2 segmentation abilities but they did not test the ability of L2 English learners with backgrounds in languages with contrastive aspiration. The current study tested native speakers of Mandarin Chinese, a language with a two-way aspiration contrast but few consonant clusters. Because these speakers use aspiration regularly in their L1 to distinguish minimal pairs, it was expected that they would outperform the participants from previously tested language groups but not reach the same levels of segmentation and discrimination of a native English speaking baseline.

A group of native English speakers and a group of native Mandarin speakers participated in this study. Thirty pairs of words (ten for each place of articulation of voiceless stops: bilabial, alveolar, and velar) where one item began with a voiceless stop (e.g., *pie*) and the other began with an *s*C cluster (e.g., *spy*) were recorded in the frame "click on this…" to create the environment for both cross-boundary and word-initial *s*C clusters. Using eye tracking with a four-item visual world paradigm (e.g., Allopenna et al., 1998; Huettig et al., 2011), participants heard the recorded stimuli and responded by selecting one of four pictures on the screen (two phonologically related experimental items and two phonologically unrelated filler items) while their eye movements were being recorded. By using eye-tracking in the visual world paradigm it was possible to collect both accuracy and real-time processing data. Collection of accuracy data allowed for comparisons to the previous research, and the addition of the real-time processing data adds to the existing body of research on L2 speech activation (e.g., Marian, Spivey, & Hirsch, 2003; Weber & Cutler, 2004).

A comparison of the overall accuracy of the two groups of speakers showed that the native Mandarin speakers were significantly less accurate in identifying the correct item in the potentially ambiguous pair, $M_{\text{English}} = 0.99$, $M_{\text{Mandarin}} = 0.91$, $t(39) = 6.7$, $p < .001$ (see Figure 1). An investigation of the real-time processing compared fixation curves of target and competitor

items (see Figure 2). ANOVAs run on measures of timing and degree of activation from these fixation curves indicate that native Mandarin speakers were slower overall to fixate target items ($F(1, 39) = 8.8$, $p = .005$), showed more fixations to competitors ($F(1, 39) = 6.5$, $p = .015$), and were slower to suppress competitors ($F(1, 39) = 25.3$, $p < .001$). Despite these differences, there was no effect of cluster type (cross-boundary or word-initial) on the fixation curves, suggesting that the native Mandarin speakers were able to adapt the contrastive aspiration system from their L1 to the allophonic system found in English. The results from this study suggest that the strategies used to discriminate ambiguous pairs in a person's L1 have a strong influence on the discrimination strategies used in their L2. When the strategies from the L1 line up with the those in the L2, learners have an advantage in L2 speech segmentation.

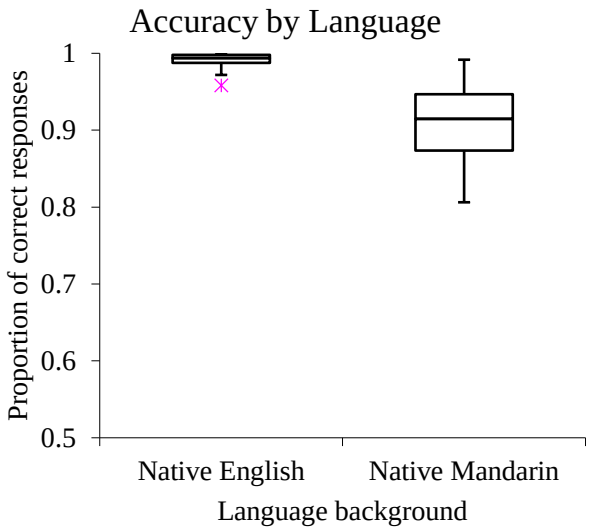

Figure 1. Accuracy by language background.

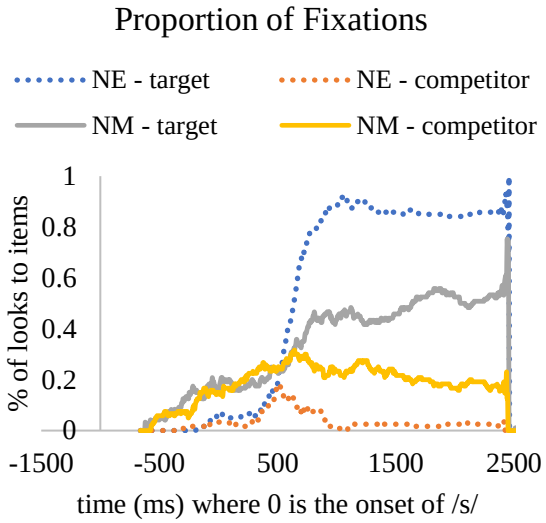

Figure 2. Proportion of fixations. NE: Native English participants. NM: Native Mandarin participants.

Selected References:

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the Time Course of Spoken Word Recognition Using Eye Movements: Evidence for Continuous Mapping Models. *JML*, *38*(4), 419–439.

Altenberg, E. P. (2005). The perception of word boundaries in a second language. *Second Language Research*, *21*(4), 325–358.

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, *137*(2), 151–171.

Ito, K., & Strange, W. (2009). Perception of allophonic cues to English word boundaries by Japanese second language learners of English. *JASA*, *125*(4), 2348–2360.

Marian, V., Spivey, M., & Hirsch, J. (2003). Shared and separate systems in bilingual language processing: Converging evidence from eyetracking and brain imaging. *Brain and Language*, *86*(1), 70–82.

Shoemaker, E. (2014). The exploitation of subphonemic acoustic detail in L2 speech segmentation. *SSLA*, *36*(4), 709–731.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *JML*, *50*(1), 1–25.