Acoustic and lexical effects on speech perception in Kaqchikel (Mayan)

Background: Kaqchikel (a Guatemalan Mayan language) has a phonemic contrast between voiceless plosives (/p t ts tʃ k q/) and 'glottalized' plosives at corresponding places of articulation (/ɓ tˀ tsˀ tʃˀ kˀ qˀ/). Our study investigates perceptual similarity within the plosives of Kaqchikel, focusing on how acoustic similarity and lexical factors affect consonant identification. Our primary results are (i) prior phonetic experience affects speech perception, consistent with EXEMPLAR THEORY (e.g. Pierrehumbert 2001, Gahl & Yu 2006), and (ii) lexical factors (e.g. FUNCTIONAL LOAD, Hockett 1955, Graff 2012, Hume & Currie Hall submitted, etc.) appear to affect on-line consonant discrimination, even in the early stages of processing.

Perception study: 44 native speakers of Kaqchikel listened to pairs of [CV] or [VC] syllables over headphones (V ∈ /a i u/, C ∈ all consonants of Kaqchikel; ISI=300ms). In target pairs, both consonants were stops; filler pairs used any other combination of consonants. Upon hearing a [C_1V]~[C_2V] or [VC_1]~[VC_2] pair, participants responded 'same' or 'different' on a button box (without time pressure).

Corpora: To assess the effect of prior phonetic experience on speech perception, we conducted an acoustic analysis of stops in a corpus of spoken Kaqchikel. Sixteen native speakers recorded free narratives with a headset microphone. The corpus was transcribed by a native-speaker linguist, and a subset (~11,000 words) was phonetically annotated with forced alignment (Gorman et al. 2011, Sonderegger & Keshet 2012). Lexical measures were computed from a written corpus of Kaqchikel (~1 million words), compiled from a range of digitized texts.

Acoustic effects: The perceptual confusability of stop consonants was modelled using mixed effect logistic regressions over the 'different' trials. The acoustic similarity of the stimuli was calculated through dynamic time warping (DTW; Sakoe & Chiba 1978); unsurprisingly, this had an effect on perceptual discriminability. The acoustic similarity of stops in the spoken corpus was also determined by DTW over ~2750 stop tokens (Mielke 2012). The acoustic similarity of stops in the corpus also had an effect on discriminability, over and above the raw acoustic similarity of the stimuli themselves. This suggests that discrimination was mediated by a representation of prior phonetic experience which includes rich acoustic detail for phoneme categories, as in exemplar theories of lexical representation.

Lexical effects: The functional load of each stop contrast was calculated as the number of minimal pairs weighted by their token frequencies (Surendran & Niyogi 2003). We found a positive correlation between the functional load of a given contrast and its discriminability in perception. Notably, this correlation *only* obtains when a measure of contextual predictability (Jeffreys divergence) is taken into account. This finding owes to the fact that /tˀ/ was highly discriminable despite having an extremely low frequency and functional load (and thus little contextual overlap with other stops). Furthermore, we found that these lexical effects persist even at very short response times (<400ms). This suggests that lexical information may influence speech perception earlier in the time course of processing than typically assumed (e.g. Kingston 2005, Babel & Johnson 2010).