Cross-linguistic trends in speech errors: Sub-lexical errors in Cantonese, John Alderete (SFU)

Overview: Very few speech error collections have baselines large enough to establish significant error patterns, and only a small handful of these analyze facts in non-Indo-European languages. As a result, these languages do not fully inform theoretical models, despite extensive research showing the importance of these under-studied languages to language production research (Chen 2000; Griffin & Crew 2012; O'Seaghdha et al. 2010). Our goal is to address this problem by documenting speech error patterns from a large collection of speech errors in Cantonese, the first for this under-studied language. In particular, we investigate Cantonese speech errors with the aim of (1) testing broad claims of phonologic encoding in a typologically distinct language, and (2) analyzing new patterns that emerge from the unique linguistic structures of Cantonese.

Methods: The data come from the Simon Fraser University Speech Error Database - Cantonese 1.0 (Alderete & Chan 2018), a database of speech errors collected from natural conversations. The speech errors were collected by four data collectors who had undergone a month-long training regime designed to master transcription skills, Cantonese linguistic analysis, and criteria for positively identifying errors. Each recording was examined by two listeners, and the errors submitted were vetted by a data analyst to ensure data quality. The errors were then classified by a psycholinguist trained in standard techniques of error classification (Stemberger 1993).

Results and discussion: Many of the Cantonese patterns reinforce known effects in the phonological encoding of segments. Sound errors in Cantonese exhibit a strong bias for single phonemes: 974 of 1159 (84%) sound substitutions involve either a single consonant or vowel, confirming to cross-linguistic trends (Nooteboom 1969). Consonant confusions also obey the phonological similarity effect (Cutler 1980; Dell & Reich 1981). In obstruent consonants, for example, 260 of 329 (79%) substitution errors change just one phonological feature, which is far above chance levels. Consonant substitutions also exhibit an output bias for frequent segments (Dell 1986), because the two most common intruders /d dz/ are high frequency, and possibly aberrant patterns arise from distinct markedness effects (Goldrick & Daland 2009).

Speech errors in Cantonese also reinforce cross-linguistic trends showing a role for syllable structure in shaping errors. The majority of sound errors that are not single phonemes (175 of 185, or 15% of the 1159 reported above) can be analyzed as coherent sub-syllabic unit, i.e., either an onset or rime, as found in Dutch (Nooteboom 1969). Sound errors also obey the syllable position effect (Boomer & Laver 1968; Fromkin 1971). That is, intruder sounds tend to have the same syllabic roles as they do in their source words. 93% of 619 contextual substitution errors retain the same roles as their sources, as in onset intruders matching onset source sounds. Furthermore, of the remaining errors, 6% of the error sounds can be analyzed as extensions of source roles, e.g., a monophthong appearing as an intruder in the first part of a diphthong.

Our Cantonese data also support roles for syllable structure in phonological encoding that have been suggested by studies of Asian languages. Cantonese diphthongs are formed by adding a high vowel /i u/ to a monophthong to form a bimoraic syllable. 27% of vowel errors involve substitutions of sub-parts of these diphthongs, and rare cases of consonants subbing for vowels generally occur in the second position of a bimoraic syllable, as found in Japanese (Kubozono 1989). Interestingly, substitutions in V1 position outnumber those in V2 by 7-to-1, indicating that serial order matters at this level. Finally, Cantonese seems to exhibit a greater than chance rate of whole syllable substitution, as found in Mandarin (Chen 2000). These empirical contributions raise new questions about how syllables and moraic structure are referenced in phonological encoding and therefore validate our research “at the margins”.

References

Alderete, John & Queenie Chan. 2018. Simon Fraser University Speech Error Database - Cantonese (SFUSED Cantonese 1.0) Burnaby, British Columbia, Canada.