When intonation fails to phonologize: the case of Southern Vietnamese
Marc Brunelle, University of Ottawa, marc.brunelle@uottawa.ca

Vietnamese is a language with a “dense” tone specification. Not only does it have a sizeable tone inventory (from 4 to 6 complex lexical tones, depending on the dialect), but a contrastive tone is realized on each of its syllables. For this reason, it has attracted a lot of attention from researchers trying to tackle the old problem of the simultaneous realization of tone and intonation, two phonological properties that should in theory compete for the same phonetic cue, f0 (Chao, 1933).

The investigation of Vietnamese intonation has largely focused on the phonetic properties that distinguish communicative and expressive functions. While some work has argued that it is the overall pitch register of sentences that matters (Trần, 1967; Hoàng, 1985), most research concludes that intonation is realized through a combination of pitch, intensity, voice quality and duration cues (Đỗ et al., 1998; Nguyễn and Boulakia, 1999; Michaud, 2005; Vű et al., 2006). More controlled studies have recently shown that there is substantial inter-speaker variation (Brunelle et al., 2012) and that intonation on short utterances can obscure lexical tone targets, some tones being more affected than others (Hạ, 2010; Hạ and Grice, 2010; Hạ, 2012).

Rather than adopting a descriptive approach, this presentation tackles the question of the nature of intonation in spontaneous Vietnamese. Given the large amount of variation found in previous studies, is there any evidence that Vietnamese intonation is phonologized, i.e. encoded through conventionalized phonological targets shared by speakers and listeners? If there is no evidence for phonologized intonation, what is the nature of the intonational variation that is attested in Vietnamese?

In order to address these questions, a corpus of Southern Vietnamese interviews, comedy shows and spontaneous conversations was collected and annotated. The corpus contains 8 hours of data collected from 19 speakers. Southern Vietnamese was chosen because voice quality plays a limited role in its tone system, thus ensuring easier pitch-tracking and limiting the number of variables to be controlled for. In this presentation, three specific research questions are addressed:

1) Are there categorical intonational targets associated to specific sentence modes (continuatives, declaratives, YN questions, open questions, alternative questions, etc.)?
2) What are the prosodic (accentual) domains to which intonational properties are associated?
3) Does the presence of phrase-final particles affect the realization of intonation? It has been claimed that intonation is clearer in the absence of such particles (Seitz, 1986; Đỗ et al., 1998).

Preliminary results based on automatic measurements of speaker normalized duration, intensity and f0 in the corpus (20,783 intonational phrases) reveal that intonational targets are organized along gradient continua rather than categorical, and that there is significant overlap between sentence modes (Figure 1). The f0 variation that is not associated to lexical tones tends to coincide with final lengthening at the edge of a prosodic domain that corresponds to syntactic clauses and adjuncts (Figure 2), which is reminiscent of intonational phrases. However, attempts to find intonational properties associated with smaller prosodic domains (downstep, trailing tones) have been unsuccessful. Finally, phrases with final particles do no exhibit less (or more) variation than others when the tone of their particles is controlled for (Figure 1).

Put together, these results point towards a prosodic system in which the heavy use of lexical tone leaves little room for phonologized intonational targets (pitch-accent and boundary tones). Southern Vietnamese prosody thus seems mostly organized around intonational phrasing and durational effects, and its information structure and communicative functions largely rely on final particles and syntactic devices. The intonation variation found in Vietnamese, although very real, seems to reflect non-phonologized effects of the frequency, effort and production codes (Ohala, 1994; Chen et al., 2004; Gussenhoven, 2004) and are thus assumed to reflect universal trends rather than grammaticalized patterns.
Figure 1: Mean speaker-normalized f0 contours of intonational phrases, per type of intonational phrase (n > 100). Ribbons represent the 95% confidence interval of the mean.

Figure 2: Mean speaker-normalized duration of intonational phrases without final particles, per type of intonational phrase (n > 100). Ribbons represent the 95% confidence interval of the mean.

References