A corpus-based study of prominence and stress in Vietnamese
Marc Brunelle, University of Ottawa, marc.brunelle@uottawa.ca

Despite a now sizeable body of research on Vietnamese phonology and phonetics, there is considerable disagreement about the nature of stress and prominence in that language. For instance, it has been proposed that Vietnamese has no word stress (Emeneau, 1951), has morpho-syntactically-determined word stress (Thompson, 1963; 1965; Ngô, 1984; Cao, 2003 [1978]), or has iambic stress (Trần, 1967; Nguyễn and Ingram, 2006; 2007; Phạm, 2008; Nguyễn, 2010). Other authors have proposed that stress is a strictly phrasal phenomenon in Vietnamese (Thomas, 1962; Thompson, 1963; 1965; Hoàng and Hoàng, 1975; Schiering et al., 2010), an issue complicated by the absence of a clear divide line between the phrase and the word in that language (Noyer, 1998).

This absence of consensus is largely due to the fact that Vietnamese is a tone language where stress, focus and phrasal prominence, if present, are not realized with pitch-accents (Michaud and Vù, 2004; Michaud, 2005; Brunelle et al., 2012), but failure to distinguish the various types of prominence in the few controlled experiments conducted so far have made the problem worse. The current study aims at uncovering and teasing apart various linguistic and discourse factors that underlie variation in syllable duration, intensity and fundamental frequency, three acoustic properties known to play a significant role in stress and prominence, in a small transcribed and annotated corpus of spontaneous and semi-spontaneous corpus of southern Vietnamese. The 87,000-syllable corpus includes speech from 16 speakers recorded in one-on-one conversations, in TV interviews and in comedy skits.

The main factors that have been investigated are linguistic factors like word stress, focus, lexical category and pre-boundary lengthening (Beckman and Edwards, 1990; Cutler and Butterfield, 1990a; b; Turk and Shattuck-Hufnagel, 2000; Byrd and Krivokapić, 2006). Discourse factors have also been included in the analysis, such local speech rate, lexical frequency and recency effects (Jurafsky et al., 1998; Fosler-Lussier and Morgan, 1999; Jurafsky et al., 2001; Bybee, 2002; Jurafsky et al., 2002; Bell et al., 2003; Gahl, 2008; Bell et al., 2009). Results were analyzed using mixed models.

Preliminary results on the entire corpus reveal no positive evidence for word stress, but a strong pre-boundary lengthening and an apparent asymmetry between function words and lexical words suggestive of a form of cliticization. Frequency effects are strong in lexical words, where a high lexical frequency correlates with shorter duration, but this effect is much weaker in function words. Recency effects are present, but tend to be more robust in conversations than in less informal TV interviews and skits. In terms of acoustic cues, most of the prominence effects in the corpus appear to be due to variation in duration. Changes in intensity and f0 seem to be largely conditioned by durational properties.

The overall picture seems to suggest that Vietnamese stress and prominence are largely determined by phrasing, and that word-level prominence is negligible. Additional work to be completed in the next few months includes a systematic revision of the corpus, a better assessment of recency effects and improvements on automatic pitch tracking.
References


Jurafsky, D., Bell, A., and Girard, C. (2002). The role of the lemma in form variation, Laboratory Phon. 7, 3-34.


