

Final lengthening at Tone Sandhi Group boundaries in Taiwan Southern Min: Boundary strength and surprisal

Sheng-Fu Wang

Institute of Linguistics, Academia Sinica (Taiwan)
sftwang@gate.sinica.edu.tw

This study examined syllable duration as a cue to the right edge of Tone Sandhi Groups (TSGs) in Taiwan Southern Min (TSM). There are two research questions. First, what is the durational pattern before TSG breaks? Specifically, does the presence of pre-boundary lengthening at TSG breaks rely on the overlap between TSG breaks and boundaries of prosodic units such as intermediate phrases (ip) and intonational phrases (IP)?

Tone sandhi in TSM is a phonological process where all syllables except for the rightmost one in a TSG, which correspond to syntactic constituents [1] categorically switch their tones. The boundaries of TSG are defined based on the presence of syllables without a categorical tonal change, i.e., in their citation tones. For example, in /hak⁴⁺² sin⁵⁵# aj^{21→53} t^hak⁴⁺² ts^he²#/ “Students love/have to read books”, the TSGs /hak⁴⁺² sin⁵⁵/ “students” and /aj^{21→53} t^hak⁴⁺² ts^he²/ “love/have to read books” are identified based on the presence of /sin⁵⁵/ and /ts^he²/ in the citation tones. This definition is different from that of ip and IP breaks in TSM, which are defined and identified based on the gradient cues such as lengthening and pitch range [2]. For the present study, we examined whether TSG breaks, defined without reference to prosodic cues, exhibit similar pre-boundary lengthening patterns as previously found in TSM [3].

The second research is how the correlation between high surprisal and longer duration changes in different prosodic conditions. The correlation between a linguistic unit’s predictability (i.e., lexical frequency, conditional probability) and its acoustic cues has been well-established [4, 5, 6]. We aimed to test the hypothesis that the prosodic structure may constraint the direct relationship between acoustic cues and information [7, 8] by examining the correlation between syllable duration and surprisal at different prosodic conditions. The measurements include unigram surprisal $-\log P(\text{word})$ representing lexical frequencies and forward/backward bigram surprisal $-\log P(\text{word}/\text{context})$ representing local contextual probability [5, 6].

Speech data were extracted from an 8-hour spontaneous speech corpus [2] with materials from 16 speakers, which contained annotations of TSGs, ips, and IPs. The data that went into analysis contained 10934 TSGs made up of 46414 syllables (35545 words). To control for the influence of prosodic conditions, we investigated syllables at TSG boundary matching with an intonational phrase boundary (TSG+IP), an intermediate phrase boundary (TSG+ip), and neither (TSG-only). Surprisal were obtained from a language model trained with the SRILM toolkit [9] with modified Kesner-Ney smoothing [10] on a written corpus containing 4.7 Million words [11].

Results showed penultimate and final lengthening at all three types of TSG boundaries ($p < .0001$ for all pairs of comparisons). In other words, TSG boundaries exhibited pre-boundary lengthening even without overlapping with larger ip and IP breaks. The overlap mostly resulted in an incremental lengthening of the final syllable: TSG+IP boundaries have the longest final syllable, followed by TSG+ip boundaries and TSG-only boundaries ($p < .001$).

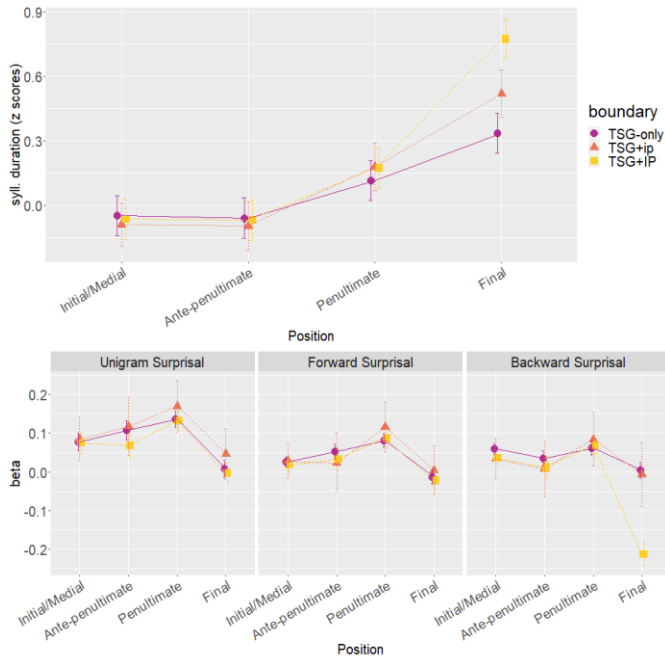


Figure 1. Upper panel: syllable duration as a function of position (x axis) & boundary type (color/shape). Lower panel: effect of surprisal as a function of position (x axis) & boundary type (color/shape)

To sum up, TSG breaks overlapping with different levels of prosodic breaks share similar pre-boundary lengthening in terms of the domain, with the presence of a larger prosodic mainly resulting in a larger size of final lengthening. As for the interaction with surprisal effects, the positive correlation between high surprisal and longer duration is found to be neutralized at the final position, which is consistent with the view that prosodic marking of boundaries may constrain predictability effects. The neutralizing trend was also found to be consistent regardless of boundary strength other than between backward surprisal and TSG+IP boundaries.

References: [1] Lin, J.-W. 1994. Lexical government and tone group formation in Xiamen Chinese. *Phonology*. [2] Wang, S.-F. & Fon J. 2013. A Taiwan Southern Min spontaneous speech corpus for discourse prosody. *Tools and Resources for the Analysis of Speech Prosody*. [3] Wang, S.-F. & Fon J. 2015. Syllable duration and discourse organization at intonational phrase boundaries in Taiwan Southern Min. *ICPhS*. [4] Jurafsky, D., Bell, A., Gregory, M. & Raymond, W. D.. 2001. Probabilistic relations between words: Evidence from reduction in lexical production. *Typological Studies in Language*. [5] Seyfarth, S. 2014. Word informativity influences acoustic duration: Effects of contextual predictability on lexical representation. *Cognition*. [6] Tang, Kevin & Jason A Shaw. 2021. Prosody leaks into the memories of words. *Cognition*. [7] Aylett, M. & A. Turk. 2004. The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech*. [8] Turk, A. 2010. Does prosodic constituency signal relative predictability? A smooth signal redundancy hypothesis. *LabPhon*. [9] Stolcke, A., Zheng J., Wang W. & Abrash A. 2011. SRILM at sixteen: Update and outlook. *IEEE automatic speech recognition and understanding workshop*. [10] Chen, S. F. & Goodman, J. 1999. An empirical study of smoothing techniques for language modeling. *Computer Speech & Language*. [11] Iunn, U.-G. 2005. Taiwanese corpus collection and corpus based syllable/word frequency counts for written Taiwanese. *ROC NSC 93-2213-E-122-001*.

As for the effect of surprisal as a function of prosodic conditions, we found that unigram surprisal was significant at all conditions except for the final position across three boundary types ($p < .05$). Forward bigram surprisal was significant for all but the final position for TSG-only breaks ($p < .001$) and only significant for the penultimate position TSG+ip and TSG+IP ($p < .001$). Finally, backward bigram surprisal had a positive effect for TSG-only breaks at all positions except for the final ($p < .01$), for TSG+IP breaks at initial/medial and penultimate positions ($p < .01$), and notably, a negative effect at the final position of TSG+IP breaks ($p < .0001$), which potentially suggest syllables in words with low likelihood at the utterance-final position were lengthened.