



Stress, tonal alignment, and phrasal position in Singapore English

James S. German¹, Adam J. Chong²

¹Aix-Marseille Univ, LPL, CNRS, Aix-en-Provence, France

²Queen Mary University of London, United Kingdom

james.german@lpl-aix.fr, a.chong@qmul.ac.uk

Abstract

Within a language, the way that f_0 marks stressed syllables can depend on phrasal position. In this study, we examine how the relationship between tonal alignment and stress in Singapore English differs as a function of position within an Intonation Phrase (IP). Singapore English speakers produced sentences containing trisyllabic words with either initial or medial stress. These occurred in Accentual Phrases (APs) that were either medial or final in the IP.

For both positions, two broad patterns emerged. In one pattern, a low trough early in the AP was followed by a rise. In the other, there was a fall from the previous AP to a low elbow followed by a plateau. For both positions and patterns, the early L aligned to either the beginning of the AP or the stressed syllable. The alignment of the peaks (rises) and fall onsets (plateaux) differed by position. For medial targets, this aligned to the end of the AP, whereas for final targets, it aligned to the end of the stressed syllable.

These findings contrast sharply with earlier findings that tonal alignment is not sensitive to stress for IP-initial words, suggesting that cues to phrasal and word-level prominence involve a complex trading relation.

Index Terms: intonation, prosody, Singapore English, stress, tonal alignment, phrasal position

1. Introduction

In a given language, the way in which f_0 is used to mark stressed syllables can differ depending on their position within the larger intonational phrase. This difference has mostly been described in terms of a difference between a nuclear pitch accent (the most prominent accent) and pre-nuclear pitch accents [1], [2], [3]. Depending on whether it is in nuclear or pre-nuclear position, the same pitch accent type (e.g., H^* , $L+H^*$) can have different phonetic realisations in terms of f_0 excursion, alignment, or both. In this study, we extend this line of investigation by examining the relationship between tonal alignment, stress, and phrasal position within an utterance in Singapore English.

In previous work, the intonational melody of Singapore English (SE) has generally been characterized as a series of rising melodies across an utterance [1, 2, 3]. [4] previously proposed that the domain of each tonal rise was a prosodic unit, an Accentual Phrase (AP), usually encompassing a single content word and any preceding function words. Further, it has often been noted that initial APs show a much larger pitch range compared to non-initial APs [4, 5, 6].

[7] further investigated the effects of stress on f_0 . They found that, unlike many other varieties of English (e.g. American English), H tones in SE generally align with the end

of the content word/edge of an AP, and not stressed syllables, with additional durational cues that mark the right edge of the AP [8]. In fact, [7] found that, at least in initial APs, there was no clear tonal target on stressed syllables, with the tonal contour shape being similar regardless of the position of stressed syllables in a content word. Words with an initial stressed syllable showed a higher overall scaling of f_0 over the entire phrase, compared to words with non-initial stress. This accords with [9]'s observation that f_0 is not a reliable cue for lexical stress. More generally, it has also previously been noted native listeners have difficulty distinguishing between stressed and unstressed syllables in SE [5, 9].

Given the difference in the f_0 range between initial and non-initial APs, we extend [7] and [8] by examining the tonal melodies in non-initial APs. Specifically, we investigate whether tonal alignment in non-initial position is sensitive to the locations of lexically stressed syllables. In doing so, we further hope to shed light on the discussion regarding the use of f_0 to cue lexical stress, taking into account how this could vary depending on phrasal position.

2. Methods

2.1. Participants

Eight native speakers of Singapore English (4 male, 4 female; mean age: 22) participated in the study as paid volunteers. All were undergraduate or graduate students at Nanyang Technological University at the time of the study. [10] reports that intonation patterns differ across the major ethnic groups of SgE speakers. To minimize possible sources of variation, only participants who self-reported being of Chinese ethnic heritage were included in the study.

2.2. Materials

The materials consisted of sentences involving trisyllabic target words with either initial or medial stress. These sentences were manipulated in two ways. First, the position of the AP containing the target words (medial or final) was manipulated through the inclusion or exclusion of a short prepositional phrase at the end of the sentence. The location of the target word with respect to the left edge of the AP was manipulated by the inclusion or exclusion of a preceding determiner or auxiliary verb. Target words with initial and medial stress were paired based on overall segmental similarity, and the two members of each pair occurred in identical sentence contexts. (1) and (2) show all four sentential contexts for the initial and medial stress members, respectively, of one pair of target words, with a full set of stimuli shown in Table 1.

12 target words (six pairs) therefore occurred in four sentential contexts each for a total of 48 items. The 48

experimental items for the present study were combined into a single list with 73 additional items from a separate experiment.

- (1) Initial stress: *minerals*
 - a. Medial AP: They explain *minerals* to the tourists
 - b. Medial AP + Function word: They explain *the minerals* to the tourists.
 - c. Final AP: They explain *minerals*.
 - d. Final AP + Function word: They explain *the minerals*.
- (2) Medial stress: *memorials*
 - a. Medial AP: They explain *memorials* to the tourists
 - b. Medial AP + Function word: They explain *the memorials* to the tourists.
 - c. Final AP: They explain *memorials*.
 - d. Final AP + Function word: They explain *the memorials*.

Target sentences	
1.	They clean (the) animals (in the back).
2.	They clean (the) enamels (in the back).
3.	They explain (the) minerals (to the tourists).
4.	They explain (the) memorials (to the tourists).
5.	They compare (the) seminars (for the report).
6.	They compare (the) tsunamis (for the report).
7.	The officers (will) mobilise (the troops).
8.	The officers (will) manoeuvre (the troops).
9.	The players (will) dominate (their rivals).
10.	The workers (will) demolish (their rivals).
11.	The soldiers (will) liberate (the village).
12.	The soldiers (will) remember (the village).

Table 1: Target sentences. Parentheses indicate elements that were manipulated.

2.3. Procedure

Participants were recorded in a sound-attenuating booth at Nanyang Technological University in Singapore. Sentences were presented one at a time as text on a computer screen. Experigen [11] was used to control presentation, which was self-paced and randomized. Participants were instructed to read sentences out loud using a natural, conversational style. Their productions were recorded at a 48kHz sampling rate with 16-bit encoding using a Shure SM81 microphone linked to a FocusRite Saffire PRO40 audio interface. Participants read the entire list two times, with a short break in between.

A preliminary investigation of our data revealed that speakers produce a variety of global intonation patterns even for the same items. Averaging of phonetic measures over qualitatively different outcomes for any given item or case would obscure this variation, and potentially lead to inappropriate generalizations. For the present study, we therefore use a more abstract, phonological coding procedure in order to explore how differences in stress lead to broad differences in tonal alignment and scaling for each prosodic position, as well as to explore the variation that can result in

each case. To do this we annotated key tonal events in each target word using the model of SgE intonational phonology developed in [4, 7, 8]. This included identifying the syllable-level alignment of L and H in each target AP, as well as noting whether the AP-level tune was rising, rising-falling, or level.

3. Results

Since productions were more fluent overall in the second block, we chose to only analyse data from this block. After excluding disfluent productions, in which speakers produced prosodic breaks within an IP, we had a total of 361 sentence productions in our dataset. Phonetic segmentation and labeling were first carried out automatically using the SPPAS force alignment tool [12]. The tonal targets for each target word were then hand-labeled and analysed in Praat [13].

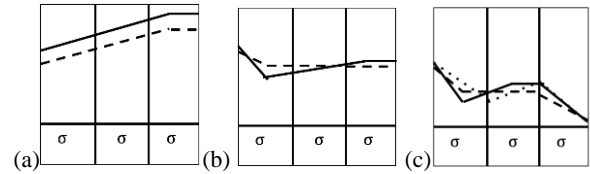


Figure 1: Tonal schemas: (a) IP-initial – solid: initial stress; dashed: medial stress; (b) IP-medial – solid: rising (initial & medial stress); dashed: plateau (initial & medial stress); (c) IP-final – solid: rising (initial stress); dotted: rising (medial stress); dashed: plateau (initial & medial stress).

In both IP-medial and IP-final position, the tonal melodies differed significantly from what has previously been observed in IP-initial APs. Whereas in IP-initial APs, [7] did not observe any clear evidence of tonal targets on stressed syllables (Figure 1a), in non-initial APs, f_0 showed sensitivity to a stressed syllable much more often. Tonal patterns across all positions in an IP are schematised in Figure 1 and a full count of different melody types is shown in Table 2. We describe the tonal patterns in each position in turn below.

IP-medial APs. In IP-medial APs, two major patterns of tonal alignment were observed. The most common tonal melody was a rising contour across the entire AP (solid line in Figure 1b). A low f_0 valley or elbow (L) occurs near the left edge of the AP, and is then followed by a rise to a high tonal target (H) towards the end of the final syllable of the AP. The AP-final H target is similar to what [7] observed in initial APs. The AP-initial L tone, however, varies in its alignment. It most commonly aligns to the stressed syllable of the content word (Figure 2), but often also aligns just to the left edge of the AP (either a preceding function word or unstressed syllable of the target word).

Melody	L alignment	Phrasal position	
		Medial	Final
Rise (Rise-Fall)	Stressed σ	31% (56)	7% (12)
	Left	14% (26)	14% (26)
Plateau	Stressed σ	19% (35)	28% (51)
	Left	12% (22)	19% (35)
Fall	-	-	22% (39)
Other	-	23% (42)	9% (17)
Total	-	181	180

Table 2: Summary of frequency of tonal patterns by phrasal position

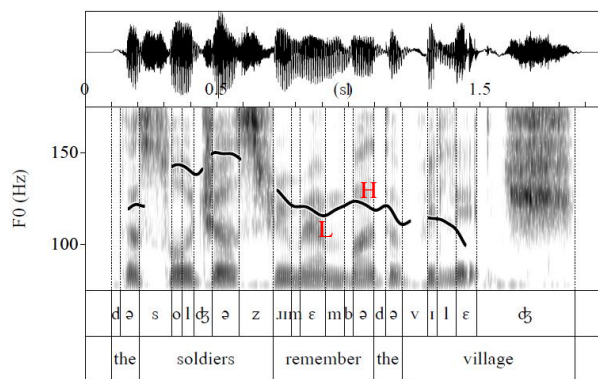


Figure 2. L tone aligned to stressed syllable followed by small rise on final syllable in an IP-medial target.

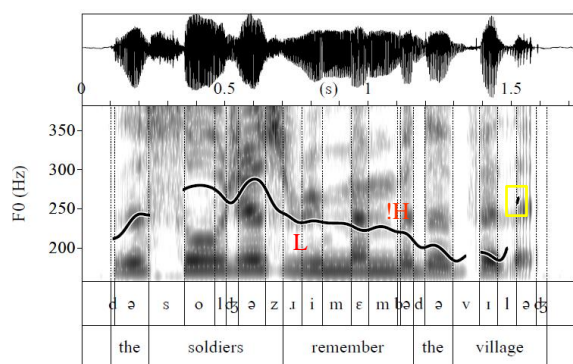


Figure 3. L tone aligned to AP-initial syllable followed by tonal plateau in an IP-medial target. (Yellow indicates pitch tracking error.)

A second major pattern in IP-medial APs involves a fall in f0 from the H tone of the previous (initial) AP to a high plateau that spans the entirety of the medial AP (Figure 3). As with the rising contours, the landing site of f0 from the previous AP (L elbow) varies between a lexically stressed syllable and the left edge of the AP (unstressed syllable). In the case of the example of Figure 3, the L elbow aligns to the unstressed syllable of *remember*. This is in contrast to the same target word in Figure 2, where the L valley is associated to the stressed syllable. Provisionally, we analyse the plateau pattern as leftward spreading of the AP-final H tone to the landing site of the AP-initial L. [4] had previously observed this pattern and had suggested that in spite of these plateaux having approximately level pitch, this nevertheless gives rise to the percept of a rise given that the level f0 goes against general declination in f0 across the utterance.

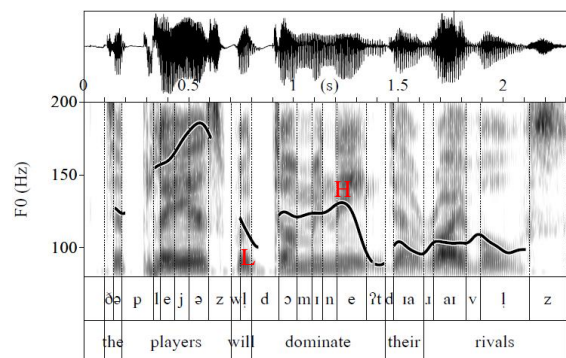
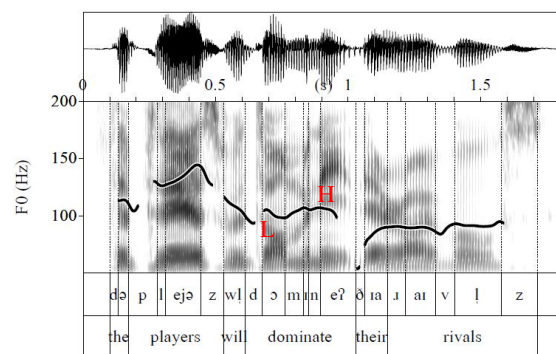


Figure 4. Top: L aligned to stressed syllable on *dominate*; Bottom: L aligned to unstressed syllable of *will*.

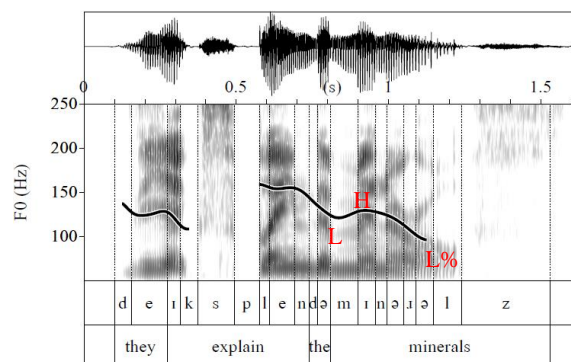


Figure 5. Final rise-fall (*minerals*) with H on stressed syllable and L tone aligned to onset of stressed syllable.

In Figure 4, we see further evidence that the AP-initial L tone can variably align to either a stressed syllable (top: first syllable of *dominate*) or to a preceding unstressed syllable at the left edge of the AP (bottom: on *will*).

IP-final APs. A similar pattern of tonal melodies was also found in IP-final APs. As with medial APs, AP-initial L tones variably align to the left edge of the AP or seek out the stressed syllable of the content word. The AP-final H tone also moves leftward, likely to allow for the realisation of the intonational phrase-final L% on the final syllable. Interestingly, this H tone also typically seeks out a stressed syllable, as can be seen in Figure 5, where the AP-final H tone aligns to the vowel of the stressed syllable in *minerals*.

In the second, more common, pattern, the final H is realised as a plateau that stretches from an L elbow near the left edge of the final AP to a fall on the final syllable of the content word (Figure 6). In Figure 6, the L elbow aligns with

the left edge of the content word on an unstressed syllable. We similarly analyse these plateaus as leftward spreading of the final H tone to the landing site of the AP-initial L. In most cases, the L% boundary tone results in a sharp fall in the final syllable which is often accompanied by strong laryngealisation.

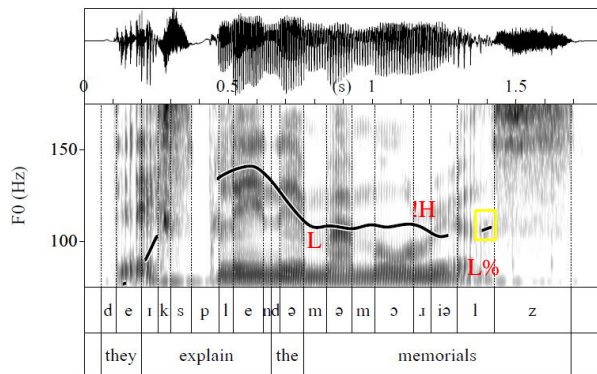


Figure 6. Final plateau on *memorials* and final L% with accompanying laryngealisation. (Yellow indicates pitch tracking error.)

The final tonal pattern we found in IP-final APs involves an f0 fall that spans the entire AP, starting from the first syllable (Figure 7). This occurred with target words with both initial and medial stress, and even when a preceding function word was present.

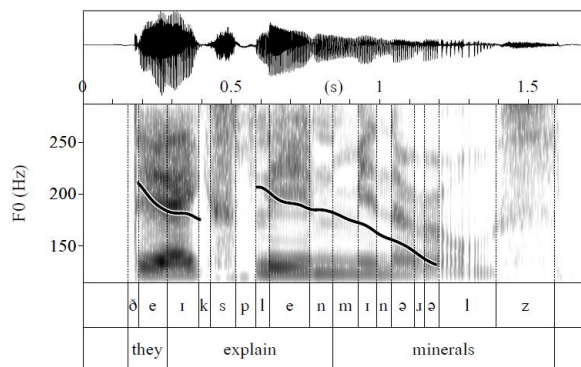


Figure 7. Fall spanning entire AP (*minerals*).

4. Discussion

In this paper, we investigated the relationship between stress, phrasal position and tonal alignment in Singapore English. Previous work by [9, 14] had only examined the relationship between lexical stress and f0 in utterance initial position. In our current study, we extended this investigation to non-initial positions. We found that, though variable, the sensitivity of f0 movement to stressed syllable position in both IP-medial and IP-final APs contrasts significantly with the lack of such sensitivity in IP-initial APs [5]. Instead, in both medial and final APs, the tonal targets often align to stressed syllables, though the details differ even between the two positions.

In medial APs, we observed two major types of patterns: either a rising contour, as in IP-initial position, or a tonal plateau. In both cases, the AP-initial L tone can variably align to the stressed syllable of the content word or to the left edge of the AP to an unstressed syllable. We found a similar pattern in final APs. Interestingly, the AP-final H tone in IP-final APs often moves leftward seeking out a stressed syllable. In cases where the initial L of the same AP also seeks out the stressed syllable, the stressed syllable bears both AP tones, such that stress is essentially marked through two separate mechanisms (i.e., attraction of the initial L and attraction of the final H).

The results of our study support the assertion that lexical prominence is cued differently via f0 depending on phrasal position. [7] previously found that in IP-initial position stress was not cued by f0 valleys or peaks, but by the overall scaling of f0 across the entire AP. However, the results of the current study provide evidence that f0 can sometimes be used to mark stressed syllables through an f0 valley or elbow (L tone). Our results do not support [14]'s model in which stressed syllables are assigned a mid tone. The variability of f0 alignment to stressed syllables might also explain [9]'s observation that SE listeners do not rely on f0 for stress perception. If future perception studies reveal that stress is equally recoverable across different positions in an IP, then this would suggest that cues to prominence in Singapore English involve a somewhat complex trading relation.

The variability in L tone alignment at the left edge of the AP also has implications for an intonational phonological model of SE intonation. Given that the L tone does not always align with a stressed syllable, this would argue against an analysis of the L tone as a L* pitch accent, since pitch accents are usually consistently associated with a prominent syllable. [4] had previously also suggested two L tones: one an AP-initial L that demarcates the left edge of the AP, and an optional L* that only variably appears on stressed syllables. A revision to this analysis, bearing parsimony in mind, might argue instead for a single L tone that demarcates the left edge of an AP, but by reassociation, restructuring, or perhaps phonetic implementation, can variably align to a stressed syllable. Future work will examine whether these different contours show differences in meaning, and will shed light on whether they should be analysed in terms of different phonetic implementations of a single underlying tonal melody as in Korean [15, 16] or belong to different phonological categories.

5. Acknowledgements

This study was made possible through support by the A*MIDEX project (n° ANR-11-IDEX-0001-02) funded by the Investissements d'Avenir French Government program, managed by the French National Research Agency (ANR), by a grant from the Singapore Ministry of Education Academic Research Fund Tier 1 (2013-T1-002-169), and by research funds from the UCLA Department of Linguistics.

6. References

- [1] K. Silverman, and J. Pierrehumbert, J. "The timing of prenuclear high accents in English," in Kingston, J. & Beckman, M. (eds.), *Papers in Laboratory Phonology I: between the grammar and physics of speech*. Cambridge: CUP, 1990.
- [2] D. R., Ladd, I. Mennen, A. Schepman. "Phonological conditioning of peak alignment of rising pitch accents in Dutch," *J. of the Acoustical Society of America*, 107, 2685-2696, 2000.

- [3] D. Mücke, M. Grice, J. Becker, and A. Hermes. "Sources of variation in tonal alignment: Evidence from acoustic and kinematic data," *J. of Phonetics*, 37, 321-338, 2009.
- [4] A. J. Chong. "Towards a model of Singaporean English intonational phonology," in *Proceedings of the Meeting on Acoustics 19*, 2013, pp. 1-9.
- [5] D. Deterding, "The intonation of Singapore English," *J Int Phonet Assoc* 24, pp. 61-72, 1994.
- [6] E.-L. Low "Is lexical stress placement different in Singapore English and British English?" in A. Brown, D. Deterding, E.-L. , Low (eds), *The English Language in Singapore: Research on Pronunciation*. Singapore, Singapore Association for Applied Linguistics, pp. 22-34, 2000.
- [7] A. J. Chong, and J. S. German, "Prosodic phrasing and F0 in Singapore English," in *Proceedings of the 18th International Congress of Phonetic Sciences* (Glasgow, UK), 2015.
- [8] A. J. Chong, and J. S. German, "The accentual phrase in Singapore English," *Phonetica*, 74(2), pp. 63-80, 2017.
- [9] Y. Y. Tan, "Is the stressed syllable stressed? The perception of prominence in Singapore English," in A. Hashim, N., Hassan (eds), *Varieties of English in Southeast Asia and Beyond*. Kuala Lumpur, University of Malaya Press, pp. 113-152, 2006.
- [10] Y. Y. Tan, "Singing the same tune? Prosodic norming in bilingual Singaporeans," in M. C. Ferreira (ed), *Multilingual Norms*. Frankfurt, Lang, pp. 173-194, 2010.
- [11] M. Becker, and J. Levine, Experigen – an online experiment platform, 2013. Available online at: <http://becker.phonologist.org/experigen>.
- [12] B. Bigi, "SPPAS: multi-lingual approaches to the automatic annotation of speech," *Phonetician 111-112*, pp. 54-69, 2015.
- [13] P. Boersma, and D. Weenink,. *Praat: doing phonetics by computer*, 2015. Computer program.
- [14] E.-C. Ng, "Reconciling stress and tone in Singaporean English," in L. Zhang, R Rubdy, L. Alsagoff (eds.), *Asian Englishes: Changing Perspectives in a Globalised World*. Singapore, Pearson Longman, pp 48-59, 2011.
- [15] S.-A. Jun, "KToBI (Korean ToBI) Labeling Conventions: Version 3.1", *UCLA Working Papers in Phonetics*, 99, pp.149-173, 2000.
- [16] S.-A. Jun, "Korean Intonational Phonology and Prosodic Transcription," in S.-A. Jun (ed.), *Prosodic Typology – The Phonology of intonation and phrasing*. OUP, pp. 201-229, 2005.