Towards a model of Singaporean English intonational phonology

Adam J. Chong*

*Corresponding author’s address: Linguistics, UCLA, Los Angeles, CA 90095, adamjchong@gmail.com

Singaporean English (SgE) is spoken in Singapore, a nation of roughly 4 million people in Southeast Asia. Recent research has sought to identify the systematic features that make it distinct from other varieties of English. This is true of prosody as well. While the intonation of SgE has been described previously within a phonetic framework (Deterding 1994; Lim 2004), no phonological model has yet been proposed. This paper proposes a model of intonational phonology for SgE within the Autosegmental-Metrical framework. Three speakers were recorded reading declarative and question sentences of varying length. Preliminary results suggest that SgE has three prosodic units above the word: the Accentual Phrase (AP), Intermediate Phrase (ip) and Intonational Phrase (IP). An AP is slightly larger than a word and is defined by a L* pitch accent on a stressed syllable and a High boundary tone (Ha) on the AP-final syllable. In monosyllabic APs we see a rising contour, suggesting that both the L* and Ha targets are met on the same syllable. The initial AP shows a larger pitch range, and subsequent APs within the same ip show a compressed pitch range and iterative downstep. Tones of larger prosodic units will also be discussed.
INTRODUCTION

Singaporean English (SgE) has been argued to have emerged as a nativized variety of English due to contact with languages such as Malay and Southern Chinese languages such as Hokkien (Platt & Weber 1980, Ho & Platt 1993). Recent research has acknowledged the fact that SgE has systematic features that set it apart from other varieties of English (see Lim 2004). This is also true of the prosody of SgE which displays a number of characteristics that make it distinct from Standard British English (e.g. Low et al. 2000). While the intonation of SgE has been described previously (Deterding 1994, Lim 2004, Ng 2011), no phonological model has yet been proposed. This paper, thus, proposes an intonational model within the Autosegmental-Metrical framework (Pierrehumbert 1980, Beckman & Pierrehumbert 1986) of intonational phonology for SgE.

SINGAPOREAN ENGLISH

English is the official language of instruction and government in Singapore. Additionally, children also learn one of Mandarin Chinese, Malay or Tamil in school depending on their ethnic background. As a result, most of the population is bilingual to varying degrees. Like many other varieties of English, SgE displays much variation although it is unclear how best to characterize this. Gupta (1998) adopts a diglossic model in which Colloquial SgE is the L(ow)-form used in casual situations and Standard SgE is the H(igh)-form used in more formal situations. Others such as Pakir (1991) have proposed a linguistic continuum instead where Colloquial SgE and Standard SgE represent the endpoints of the continuum. In the present paper, I adopt a diglossic model (Gupta 1998). SgE is defined as Standard Singaporean English, since we can reasonably assume that when reading speakers will adopt the H-form. I leave open the question whether or not the present model can be expanded to capture patterns in Colloquial SgE for future work.

A number of previous studies have investigated certain prosodic characteristics of SgE. Ng (2011) in particular has proposed a close relationship between lexical stress and tone assignment in Colloquial SgE. Specifically, she argues that the domain of tone assignment is the syllable, arguing essentially for categorical tone that is predictable from the stress placement on the word. The present study takes her analysis as a starting point, but departs from it in positing an intonational model. Due to the constraints of the present paper, however, I do not discuss the differences between Ng (2011) and the present model.

METHOD

The present study is based on a small corpus of data collected from thee native speakers of SgE: a 20-year-old female speaker, a 24-year-old female speaker and a 27-year-old male speaker. The two females reported using English most often at home, while the male speaker reported using Mandarin Chinese more often. All three speakers were university-educated. Speakers were recorded using laptop-internal microphones. They produced sentences of varying length, stress pattern and type. As much as possible, this was done in a quiet space to ensure that the pitch was tracked as well as possible. Recordings were then segmented and labelled, and pitch tracks were examined using Praat (Boersma & Weenink 2011). Pitch tracks were analyzed by locating peaks and valleys in each utterance.

MODEL

It is proposed in this paper that SgE is a head- and edge-prominence marking language. Evidence has been found for one pitch accent type: a low tone, L*. There is also evidence for three levels above the word: the Accentual Phrase (AP), the Intermediate Phrase (ip) and the Intonational Phrase (IP). For the AP, I propose an initial aL low boundary tone and a final Ha high boundary tone. Preliminary evidence is presented to argue for the existence of a phrasal constituent between an AP and IP, that is, an Intermediate Phrase (ip). In the current data, there is evidence for both a L- tone and H- tone that mark the right edge of an ip. Finally, at the Intonational Phrase (IP) level, two boundary tones, L% and H% are proposed.
Figure 1 provides a schematic of the proposed hierarchical structure of the model. APs seem to contain only one pitch accent, suggesting that the domain of pitch accents is the AP. APs contain one or more words, but each AP can generally contain only one content word, with the possibility of a number of preceding function words. Intermediate phrases (ip) contain a number of APs and can be marked by an L- or H-. Lastly, IPs contain one or more ips and are marked on the right edge with a boundary tone, L% or H%. Whether tones of higher prosodic units override those of lower prosodic units is unclear from the present data, though there is some evidence to suggest that this is not the case. This, however, is one avenue in need of future investigation.

**Accentual Phrases and Pitch Accents**

An AP in SgE generally consists of one content word and zero or more preceding function words. An AP-final H boundary tone (Ha) is realized on the final syllable of the accentual phrase. Another L tone marks the left edge of the AP (aL). Taken altogether, the tonal melody of an AP is characterized by a rising contour due to the LH tone targets across the AP. The present model also posits a single type of pitch accent in SgE: L*. This L* pitch accent, however, is occasionally optional. It is admittedly not clear whether an L* pitch accent is necessary for us to account for the patterns we see. In fact, it has been argued elsewhere (Deterding 1994) that it is hard to determine which word in a sentence is prominent. However, the L* pitch accent is required, at the very least, to account for some patterns in the present data. In this paper, L* pitch accents are only labelled in a word which has a low F0 and is perceived as prominent. They are not labelled if F0 is high or when prominence on the stressed syllable is not perceived. In cases in which the initial syllable of an AP is also prominent, it is marked with both aL and L*.
In Figure 2, we see a clear rising contour on *Amanda*. F0 then lowers in *remembered*, reaching a valley on the stressed second syllable, presumably meeting an L* target. F0 rises a little on the last syllable before falling again during *anagrams*. What we can see from this example is that the initial AP (*Amanda*) has a larger pitch range and successive APs show successively reduced pitch range. This pattern of the initial word being produced with a larger pitch range has also been observed previously, most notably by Deterding (1994) and Low & Brown (2005). Deterding (1994) suggests that this extremely high pitch corresponds to the introduction of a new topic. The rising contour of aL Ha is most evident in the initial AP since this is where there is the largest pitch range. In fact, we also see this contour in monosyllabic APs as can be seen in Figure 3.

The rise on *John*, which constitutes an AP on its own, is suggestive that all the tonal targets aL, L* and Ha are met, at least in initial position, even when there is only one syllable for all these tones to be realized. In the second AP, *wants*, there seems to be an L target initially, but we see a plateau as opposed to a rise. Here, it seems as though
the aL target was undershot and is only realized with a high peak. Undershot of the aL is not the only way in which the phonetic realization of the phonological shape of an AP can vary. Consider the following example in Figure 4.

**FIGURE 4.** Pitch track of “Amanda memorized anagrams.”

We saw a minimal pair of this sentence in Figure 2. The only difference between the two sentences is the medial verb. In the first case, remembered is stressed on the medial syllable, and in the second memorized is stressed on the initial syllable. Whereas in Figure 2 there is an actual rise phonetically on the last syllable, in the Figure 4, the contour is a flat plateau across the entire AP, before falling again on the third and final AP of the sentence. That is, it seems as though the AP-final Ha is not always realized as phonetically high, especially outside of the initial AP. The plateau contour on the medial AP in Figure 4 does not seem to be an isolated case, as we see in Figure 5. Here, I have labeled the final Ha with a question mark since F0 is lower in the final syllable than the preceding syllable. It is unclear whether or not this means that the boundary tones of the higher prosodic units, L-L%, overrode the final H.

**FIGURE 5.** Pitch track of “Amanda dramatized the story.”

Crucially, in these cases where we see a plateau in F0, F0 stays relatively level against declination. This has the result then of giving the percept of a final rise despite the lack of an actual rise in pitch phonetically. What seems to determine the presence or absence of a rise on the final syllable of the AP is the placement of the L* pitch accent. In
words such as *remembered* and *collided* the pitch accent is on the medial or penultimate syllable, whereas in the other cases, such as *dramatized* and *memorized*, stress and thus the L* pitch accent is further away from the final syllable. The rise in these cases seems to be a function of wanting to ensure the perception of a final high signaling the AP boundary. When the pitch accent occurs earlier in the AP, a plateau is maintained for longer against declination which is enough to give the percept of a final high. When the L* pitch accent is realized adjacent to the final syllable, on the other hand, a final rise is necessary to actually cue the AP-boundary. Whether or not this generalization extends to other longer words with similar stress patterns remains to be seen. If there is free variation in the surface realization of the LH contour, the question will then be whether or not an L* pitch accent is actually necessary at all in the model.

**Intermediate Phrases**

Intermediate phrases in SgE contain a number of APs. Consider the following sentence in Figure 6.

![Pitch track of “Oliver will marry Maria and Jonathan will marry Clara.”](attachment:image.png)

**FIGURE 6.** Pitch track of “Oliver will marry Maria and Jonathan will marry Clara.”

The left edge of an ip can be characterized by pitch reset. The pitch of the ip-initial AP is not lowered in comparison with the previous AP as we would expect, nor is there a reduced pitch range in comparison. We can point to two ip-final tone types: H- and L-. H- tones are realized either as a flat plateau continuation of the Ha or a further rise on the final syllable of the ip. L- tones seem to only realized IP-finally. We see a H- tone of the first of the plateau type in Figure 6 at the end of *Maria*. We see a H- tone of the rising type in Figure 7 at the end of *Amanda* (see also Figure 9 below).
An L-IP tone is evidenced at the end of declarative sentences and wh-questions as can be seen in at in IP-final position in Figure 7 above and Figure 8 below.

Intonational Phrases

Intonational Phrases (IP) are the largest tonally marked prosodic unit in SgE and are usually the size of a simple sentence or a large phrase. A boundary tone is realized on the final syllable of the IP. There seems to be evidence for two different IP-final boundary tones, L% and H%. Declarative sentences and wh-questions end in L% and yes-no questions end in H% tones. We have already seen declarative examples above with an L% tone finally. Figure 8 shows a wh-question with an L%.

Notice that there is a small rise on *American* and then a final fall. What this suggests is that the Ha of the AP is still realized together with L- and L%. This gives us potential evidence that there is no tonal override of higher prosodic constituents, unlike in other languages such as Bengali (Khan 2008). An example of a yes-no question with...
a final H% is shown in Figure 9. Here we see a rising contour on the final syllable of tomorrow. This is presumably the result of the realization of Ha, a H- and H% all on the one syllable.

![Pitch track of “Are we running tomorrow?”](image)

**FIGURE 9.** Pitch track of “Are we running tomorrow?”

### CONCLUSION AND FUTURE DIRECTIONS

The present model of SgE intonational phonology argues that SgE is a head- and edge-prominence marking language, similar to languages such as Bengali and Georgian. It is proposed that SgE possesses a prosodic structure comprising of three tonally-marked phrases above the word.

Stress is lexically-specified and content words (and sometimes function words) seem to receive a pitch accent realized as L*. Accentual Phrases typically contain one content word and any preceding function words. APs are marked tonally by an aL initially and Ha finally. A number of APs can make up a single Intermediate Phrase which are tonally marked by either a L- or H-. Finally, Intonational Phrases represent the largest tonally-marked prosodic unit and is marked at the right-edge by a L% or H%.

A number of questions remain to be answered. The first is the relationship between tones of different prosodic units. More specifically, it is unclear whether or not tones of a higher prosodic unit override those of a lower one or not. A second question is whether or not an L* truly optional and is necessary in addition to aL.

The study of the intonation of SgE is instructive beyond the immediate descriptive goal. One future avenue of research would be cross-dialectal comparisons of intonation amongst different varieties of English. Furthermore, because of the history of its genesis, SgE provides an interesting opportunity to study the results of language contact on a language's intonation system. Moreover, it has been noted elsewhere (Lim & Tan 2001) that the SgE speakers of different ethnicities cue prominence in different ways. It is plausible then that they also differ in their intonational systems. The present study only recorded Chinese SgE speakers. A further avenue of research would be to analyze data from SgE speakers of other ethnicities. Finally, the SgE lexicon contains a number of question and discourse particles that have been argued to have lexically-specified tone (see Lim 2004 for a summary). How these would fit into the present intonational model is an open question.

This paper has presented the first phonological model of SgE intonation. Of course, many questions remain and further investigation is required to refine and modify the present model.

### ACKNOWLEDGMENTS

This paper has greatly benefited from helpful comments and discussion from Sun-Ah Jun, Marc Garellek, Chad Vicenik, Jason Bishop and members of the UCLA Phonetics Seminar group. All remaining faults are my own.
REFERENCES


