Intonational Phonology in Colloquial Singaporean English*

Kelly Banciella Smeno

A thesis submitted in partial fulfillment of the requirements for the degree of Bachelor of Arts in Linguistics

Swarthmore College
December 2015

Abstract

One of the primary goals of intonational phonology is to explore how meaning is assigned to an utterance through various suprasegmental features at the word level and above. This can be modeled using Autosegmental Metrical theory which creates a hierarchical prosodic structure for an utterance. How these different levels interact with each other is dependent on the language in question (Ladd, 2008).

This thesis explores Colloquial Singaporean English (CSE), a language native to the city-state of Singapore. Standard English along with Mandarin Chinese, Malay and Tamil are recognized as the official languages of the area. This puts CSE in a stigmatized role, despite its use in almost all spheres of life, barring government policy and writing in general. The dialect differs greatly from standard forms with heavy borrowing from languages such as Mandarin, Hokkien and Tamil on phonetic, syntactic and semantic levels (Harada, 2009).

Within CSE, there is contention on how to accurately describe its intonational system due to the lack of empirical data. One model, developed by Ng posits that each syllable of the word is assigned a specific tone based on it’s place in the word (e.g. Ng, 2011). Another model, argued for by Chong, uses Autosegmental Metrical theory to break up utterances into accentual phrases (e.g. Chong, 2013).

In order to test the merit of each model, I collected novel data from two consultants focusing on target words of varying syllable length. I then compared the two models using this data. Ng’s Model failed to accurately portray most of the data collected. Chong’s Model was more successful but in need of modification. I proposed two modifications, one that would treat prefixes as their own prosodic unit (in line with the existing findings on the prosodic behavior of prefixes) and another that allowed for multiple prosodic units that determine intonational contours within a multisyllabic word greater than four syllables.

*I would like to thank my advisor, Prof. Byron Ahn for helping me through every step of the thesis process. I would also like to thank Prof. Ted Fernald for being my second faculty reader. In addition, I would like to thank Prof. Nathan Sanders for providing aid in the initial steps of this project. Lastly, this thesis would not have been possible without the contribution of my two consultants, PK and RT.
# Contents

1 Introduction .................................................. 3

2 Background .................................................. 3
   2.1 Intonational Phonology .................................. 4
   2.2 Colloquial Singaporean English ......................... 9
   2.3 Current Competing Models ................................. 11

3 Data .......................................................... 16
   3.1 Consultant Description ................................... 16
   3.2 Data Overview ............................................. 17
   3.3 Data Description ............................................ 19
      3.3.1 One and Two Syllable Words ....................... 19
      3.3.2 Three Syllable Words ................................. 24
      3.3.3 Four or More Syllable Words ...................... 26

4 Analysis ..................................................... 28
   4.1 Prefixes as Unique Prosodic Units ...................... 29
   4.2 Multisyllabic Words and the Accentual Phrase ......... 31
   4.3 Ternarity .................................................. 34
   4.4 Exceptions ................................................ 36

5 Conclusion .................................................. 37

6 Appendix ...................................................... 41

References ..................................................... 67
1 Introduction

Colloquial Singaporean English (CSE) is an important language to study due to its contact with many other languages, which has influenced it greatly. Because of this, linguists have examined the language to discover how it differs from Singaporean English in its phonology, syntax and semantics. These subjects have been studied at length (Harada, 2009; Leimgruber, 2011; Tay, 1993). However, within the field of phonology, one continuing mystery of CSE is its intonation system.

This glaring gap in the field is due, in part, to the general lack of data on the subject. In this paper, I enter the current conversation and provide the subfield some of the data it needs. In addition, I evaluate two current models on the intonational phonology of Colloquial Singaporean English, Chong and Ng’s Models, and compare them. Ng’s Model is lexically tone based while Chong’s Model relies on AM Theory. In comparing them, I found Chong’s Model to be a better predictor of the majority of the data. However, it does need a few modifications in order to be fully representative of the data. Because of this, I propose two important modifications to one of the models.

The two modifications proposed call for a re-evaluation of what accentual phrases consist of. Chong’s Model initially has APs consisting of an entire content word and its related function words. In addition to this I propose that prefixes should be considered their own APs as well. Furthermore, longer multisyllabic words have the ability to break up into multiple APs, with a preference toward ternarity, or groupings of three.

2 Background
In order to fully comprehend how intonation is used in Colloquial Singaporean English, we need to have a fundamental understanding of the field of intonational phonology and some background on Singaporean English itself. In addition, it is important to understand the current dialogue on intonation within Singaporean English. In this section, I briefly explain intonational phonology as it pertains to this paper, touch upon Singaporean English and introduce the two current competing models for intonation within Colloquial Singaporean English.

2.1 Intonational Phonology

Intonation “refers to the use of suprasegmental phonetic features to convey ‘postlexical’ or sentence-level pragmatic meaning in a linguistically structured way.” (Ladd, 2008). In order to fully understand this definition we must understand two key terms within it: suprasegmental and postlexical.

For the purposes of this paper, the term suprasegmental will be defined as “the features of fundamental frequency (F0), intensity and duration” (Ladd, 2008). In other words, we will be looking at the pitch track of the utterance, noting where it rises or falls intensely and for how long. The fact that others have defined this term differently will not be discussed within this paper.

The term postlexical is used to define the idea of meaning existing above word level. Lexical items create meaning at word level, while postlexical units contribute meaning above this level. In this thesis we will be focusing on postlexical phonological items. Intonation is used to add meaning to phrases or entire utterances. For example, the difference between the declarative sentence *I was chosen for the constitutional convention.* and the question *I*
was chosen for the constitutional convention? is intonation. In the latter, the speaker rises in tone in order to form the question (Gunlogson, 2008). Intonational features are never used to describe lexical differences in languages. However, some languages may use the phonetic basis of intonation, such as stress or tone, to differentiate lexical items. It is important to draw a distinction between intonational phonology and the use of phonetic intonational items for differentiating lexical items.

It is also important to note that intonation is linguistically significant. This can be demonstrated for in English, where pitch-accent can change the fundamental meaning of a sentence but not the fundamental meaning of a word. For example, Figure 1 and Figure 2\(^1\) both show the same sentence, only differing in intonation. In Figure 1, the pitch-accent is on the word *steal*. If we look at the figure, we see a dramatic rise in pitch track\(^2\) on the word *steal*. The blank spot at the beginning of the word is due to the voiceless fricative [s], which has no pitch. Other gaps in the pitch track may be due to Praat\(^3\) having difficulty reading the item. In addition to the dramatic rise in pitch, the duration of the word is longer than in Figure 2. Thus, we know the pitch accent, which is only associated with stressed (and thereby longer) words, is on the word *steal*. Figure 2 has a pitch accent on the word *money*. We see this on the dramatic rise in the first syllable of the word. The first syllable of the word *money* is the lexically stressed syllable, and pitch-accent in English typically find themselves on the stressed syllable of the word. In addition, the word itself is said for a longer amount of time. Because of this difference in intonation between the two example sentences, the meaning changes. In the first, the speaker is implying that they did not steal

---

\(^1\)Both sentences in Figures 1 and 2 are spoken by me.

\(^2\)The pitch track is the F0 line that marks intonation on an utterance

\(^3\)Praat is the program being used to analyze the data.
the money, but rather, did something else to the money. In Figure 2, the speaker is implying that they didn’t steal the money, but rather, they stole some other item. The meaning of the words themselves have not changed, but rather the meaning of the entire utterance.

Figure 1: Pitch-accent on steal

Figure 2: Pitch-accent on money

It is important to note that pitch-accent should not be equated with stress, as it is easy to confuse the two within English. The source of this confusion can be found in the fact that, in stress-accent languages (like English), pitch-accent is found on the stressed syllable of the prominent word. Pitch-accents can only be found where there is stress, but where stress is found is determined independently of this. Pitch-accent is found phonologically on a higher prosodic level while lexical stress is found word level. Lastly, pitch-accents are found phonetically and phonologically in languages that do not have stress (Ladd, 2008).

These pitch events can be the result of different types of intonation systems including: stress-accent, lexical tone and pitch-accent. English is just one example of a stress-accent language. In these languages, stress is used to differentiate the meanings of lexical items. For example, the verb permit and the noun permit are only distinguishable through their stress pattern. The stress is on the second syllable of the verb form and on the first syllable of the
noun form (Ladd, 2008). Pitch-accent is used to put emphasis on the prominent words, and can only be found on the stressed syllable of the word, as noted previously. This is why, in Figure 2, the pitch accent is only found on the first syllable of the word money.

Lexical tone languages use tone to differentiate lexical meaning between words. One such language is Mandarin Chinese. For example, the words hua ‘flower’ and huá ‘speech, language’ are identical in every way except for their tone. The first has a high level pitch while the latter has a falling pitch (Ladd, 2008). In this way, tone is heavily tied with the lexicon of the language rather than the postlexical intonation system.

Finally, pitch-accent languages, such as Japanese, are usually defined as employing an intonation system which “uses pitch to mark certain syllables in the speech stream” (Venditti, 2005). In other words, pitch-accents are a lexical part of the word, like tone is in Mandarin Chinese. In addition, unlike Mandarin, but similarly to English, most syllables are not associated with a pitch event. Unlike stress-accent languages, pitch-accent does not have any prominence-lending attributes. It is also important to note that pitch-accents do not occur on every word. For example Figures 3 and 4 show identical phrases differentiating only in pitch accent. Figure 3 is not accented while Figure 4 has a pitch accent on ue. We can see this clearly by the dramatic rise in the pitch track on that specific syllable. This difference in pitch accent creates two very different lexical meanings, with the first meaning ‘something to plant’ and the latter meaning ‘the ones who are starved’. (Venditti, 2005:2)

It is important to differentiate these different language types in order to understand which category Colloquial Singaporean English belongs in. One thought could be that it is a stress-accent language, as it is a variant of English. Another thought could be that it is a lexical tone language, especially since it has extensive contact with and borrowing from
Mandarin Chinese. Lastly, there is a possibility that Colloquial Singaporean English is a pitch-accent similar to Japanese because of its use of pitch-accents in its intonation system. This topic will be further discussed in Section 5.

We can model all of these different intonational systems by using Autosegmental Metrical (AM) theory. This theory proposes two major points. One, that intonation manifests itself in local events and transitions. In other words, one does not have to describe each minute transition on every syllable in intonation, but rather the key events that make it up which are distributed more sparsely in the utterance. Secondly, it argues that utterances are comprised of a hierarchical prosodic structure made up of different levels. These levels include the intonational phrase (IP), the intermediate phrase (ip) and the accentual phrase (AP). In addition, there is a word level and syllable level. We can see how these levels are structured in Figure 5.

If we read Figure 5 from bottom to top we begin with the smallest unit, the syllable.

---

4 Which levels are present is dependent on the language in question. In addition, there may be additional prosodic levels outside the ones discussed here but those are irrelevant to this thesis.
Syllables are marked with a (S) if they are stressed and a (s) if they are unstressed. Syllables build up to words. Words then make up accentual phrases (AP). An AP is made up of a single content word (Wc) and its accompanying function words (Wf) if applicable. For example, the term *the monkey* would be one AP. Above the accentual phrase is the intermediate phrase (ip) which can have one or more APs within it. Finally, the intonational phrase (IP) encompasses the entire utterance.

![Figure 5: AM Theory Hierarchical Structure](image)

How intonational targets associated with these different hierarchies affect or override each other depends on the language (Ladd, 2008). We will examine in depth one possible AM model for Colloquial Singaporean English in Section 2.3.

With this brief background on intonational phonology, we are now able to use these models to collect and analyze data from speakers.

### 2.2 Colloquial Singaporean English

Singaporean English (SgEng) is a variety of English spoken in Singapore, a city-state south of Malaysia (Harada, 2009). Although tiny, this city is ethnically and linguistically diverse. It has four official languages: Singaporean English, Mandarin Chinese, Malay and Tamil. Of these four, SgEng has the most prestige. It is used primarily in formal situations and all
written documents. However, the focus of this paper is not on SgEng but rather on its more colloquial variant, Colloquial Singaporean English (CSE).

CSE is spoken in complementary distribution to SgEng. It varies significantly from SgEng in its syntax and phonetics, due to the influences of Mandarin, Hokkien (another Chinese dialect), Malay and Tamil (Harada, 2009). These variations include but are not limited to copula deletion, article deletion and its lack of verb marking (Harada, 2009). Because of these variations, the language has been stigmatized by the government, which has pushed for its eradication; however, many of the younger generation uphold the colloquial variant as an important part of their cultural identity (Harada, 2009).

One way CSE differs from SgEng is through article deletion. We can see in example (1a) a sentence in Singaporean English. We can compare this to example (1b) with the CSE equivalent. In (1b) the article a is not pronounced with the noun phrase, while in (1a) it must be there to be grammatical. (Data (1-3) taken from Tay, 1993:27-35)

(1a) May I apply for a car license?  (SgEng)
(1b) May I apply for car license?  (CSE)

CSE also differs from SgEng through its verb marking. CSE tends to not inflect verbs for tense, and instead, relies on syntactic structures and context to derive meaning. For example, in (2a) we see a sentence in SgEng that marks the verb go with a third person singular marker. However, in the CSE equivalent (2b), go is left in the infinitive, unmarked for tense.

(2a) He always goes there every Saturday.  (SgEng)
(2b) He always go there every Saturday.  (CSE)

In addition, CSE has multiple instances of “copula deletion”, differentiating it from the standard dialect. In sentence (3a) we see the SgEng sentence, which includes the copula
## Table 1: Selected Particles of CSE

<table>
<thead>
<tr>
<th>Particle</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ah</td>
<td>tentative marker, continuation marker</td>
</tr>
<tr>
<td>hah</td>
<td>question marker</td>
</tr>
<tr>
<td>hor</td>
<td>attempts to garner support for a proposition</td>
</tr>
<tr>
<td>lah</td>
<td>mood marker, appeals for accommodation</td>
</tr>
<tr>
<td>leh</td>
<td>marks a tentative suggestion/request</td>
</tr>
<tr>
<td>lor</td>
<td>indicates obviousness or resignation</td>
</tr>
<tr>
<td>mah</td>
<td>marks information as obvious</td>
</tr>
<tr>
<td>what/wot</td>
<td>marks obviousness and contradiction</td>
</tr>
<tr>
<td>meh</td>
<td>indicates scepticism</td>
</tr>
<tr>
<td>ya</td>
<td>conveys (weak) emphasis and uncontroversiality</td>
</tr>
</tbody>
</table>

to be in order to be grammatical. In the CSE version in example (3b), this verb is not pronounced and the sentence remains grammatical.

(3a) My handwriting is not clear. *(SgEng)*
(3b) My handwriting not clear. *(CSE)*

Furthermore, CSE varies with its vocabulary and heavy borrowing from other nearby languages. For example, CSE uses a number of clause final discourse particles, most likely borrowed from Hokkien or Cantonese, whose exact meanings are still being disputed (Leimgruber, 2011). In Table 1 we can see a number of these particles as well as their corresponding meaning according to Leimgruber. Table 1 has been reproduced exactly from Leimgruber (Leimgruber, 2011:9).

The intonational phonology of Colloquial Singaporean English has not been studied to the extent as many other parts of the language. The purpose of this paper is to explore this topic further.

### 2.3 Current Competing Models

The current dialogue on Colloquial Singaporean English’s intonation system is limited. In part due to the relatively small number of research publications on the topic, there is in-
sufficient data to make all the necessary observations in order to establish a more complete
model of the language’s intonation. There are two competing models that attempt to explain
CSE’s intonation system, one is based on the influences of Mandarin Chinese lexical tone,
and the other on AM Theory.

The former has been developed in several works by E-Ching Ng (Ng, 2011; Ng, 2012). Ng’s Model is based on the idea that CSE is a lexical tone language, like Mandarin Chinese. The model states that every syllable of a word carries a tone based on its placement within the word.

Ng presents her tone assignment generalizations as follows (Ng, 2012)⁵:

- H is assigned to the final syllable of the prosodic word.
- L is assigned to initial unstressed syllables.
- M is assigned to all remaining syllables.

She also includes examples of tone assignment with multiple words of varying syllable length. This table is reproduced in Table 2⁶ (Ng, 2012:87). Examples (a)-(e) demonstrate words with initial stressed syllables and therefore no low tones, while examples (f)-(j) show words with initial unstressed syllables. Although Ng provides this table, she provides no pitch tracks for words longer than three syllables, nor does she explicitly discuss them.

In addition, these multisyllabic examples are only measured after they are produced in isolation rather than in a sentence. This is problematic because words spoken in isolation may behave differently than when spoken in a full utterance. As discussed in Section 2.1, standard assumptions of intonational phonology include that there are multiple hierarchies encoding meaning into a phrase and these hierarchies may affect each other. If a word is

---

⁵Where H stands for a high tone, L for a low tone and M for a mid tone.

⁶It is important to note that this table has been reproduced exactly from Ng’s paper, including the ways in which stress is marked within it.
spoken in isolation, it is spoken utterance finally, and may be affected by multiple hierarchies including the IP, IP and AP. On the other hand, if the word is spoken in the middle of an utterance, it will not be affected by the IP or even the IP.

Ng's model is based on a rather limited set of CSE data. One of the few presentations of a pitch track of a sentence is reproduced in Figure 6 (Ng, 2011:36). Each syllable is marked by a tone according to the generalizations discussed previously. The word *cannot* starts with a low tone on the first syllable and ends with a high tone on the last syllable. The word *minimum* has the first two syllables notated as mid tones and ending with the last syllable as a high tone. Thus, we clearly see a low to high or mid to high pattern on each word. This is predicted accurately with her model.

![Figure 6: Example Sentence Transcribed by Ng](image)

Although Ng never mentions function words in her analysis, she marks the function words in her examples as having a low tone. We see this on the word *from* in Figure 6. I will follow suit when I test her model in Section 3.3.

---

1 It is interesting to note that *cannot* is treated as one prosodic word. It would be interesting to investigate if this were true for other words such as *will not*. Unfortunately, this topic is outside the scope of this thesis.
One problematic aspect of Ng’s analysis is her use of the term *lexical tone*. The phenomenon she describes within her paper, of syllables having an assigned tone based on location, does not mirror what one would find within a lexical tone language (in fact, this sounds more like a pitch-accent language). In particular, there does not seem to be any minimal pairs of CSE that differ only by tone. This makes it highly unlikely that the language is a lexical tone language.

In contrast to Ng’s lexical tone model, Chong uses AM theory to devise a model (Chong, 2012; Chong, 2013; Chong & German, 2015). He proposes that CSE’s intonation pattern is the result of tone targets associated with higher prosodic levels than the word level. Although some of the work is co-authored, I will be referring to this model as Chong’s Model.

Chong’s Model proposes that the low to high pattern Ng has observed is not from the lexical properties of the word and its syllables but rather from the Accentual Phrase (AP). As mentioned in Section 2.1, the AP consists of either one word, or a word and its associated function words. In this way, he specifically accounts for function words within his model, unlike Ng. Chong’s Model is reproduced in Figure 7 (Chong, 2013:2)

Chong also provides for more levels above the AP, including the Intermediate Phrase (ip) and the Intonational Phrase (IP). The boundaries for the AP are defined with the following notation: an initial low tone (aL) and an ending high tone (Ha). Both the ip and IP are edge marked by either a low or high tone. They can be differentiated in the notation with either a dash (-) or percentage mark (%) respectively. For example, a low IP-edge tone would be notated as L% while a low ip-edge tone would be marked as L-. (Chong, 2013)

Chong’s Model also differs in that it allows for a low tone pitch accent (L*) that can be found within an AP. The inclusion of a low tone pitch accent (L*) creates a key difference
in the two models. For one, the L* can only occur on a stressed syllable. In Ng’s model, it is impossible for a stressed syllable to have a low tone. However, it will become apparent in the data in Section 3.3 that words with a stressed low tone do, in fact, exist. An example of one can be taken from Chong’s own paper, reproduced in Figure 8 (Chong, 2013:3).

In Figure 8 we see that the utterance is comprised of three APs framed by an aL and Ha notation. Each has a low pitch accent (L*) on a syllable of the word. The utterance ends with a low ip (L-) and a low IP (L%). In this way we see how both word boundary and edge boundary (ip and IP) can affect the pitch of an utterance.

Because of the relatively low amount of research in this area, Chong’s analysis (like Ng’s) is built upon a small empirical base, and it must be tested on a larger data set. One major gap in all previous work is systematic investigation of multisyllabic words greater than three syllables. Ng touches upon the issue very briefly in an example, but presents the words in isolation rather than in a sentence (Ng, 2012). This is a gap I address with an experimental design that manipulates syllable length as a variable.
Figure 3: Example Sentence Transcribed by Chong

3 Data

3.1 Consultant Description

The data used in this research was collected from two consultants. They will be referred to as PK and RT in this paper. To preserve anonymity, these initials were formed by combining a random pair of letters. Both consultants are native Singaporean English speakers.

PK is fluent in Mandarin Chinese, is adequate in Malay and Hokkien, and has some knowledge of Japanese and Cantonese. She is 22 years old, born and raised in Penang, Malaysia. However, she spent her Middle and High school years in boarding school in Singapore. She is ethnically Straits-born Chinese.

RT was born in Singapore and spent his life until university in the city state. He has an
adequate grasp of Chinese and some knowledge of Malay, Indonesian and Cantonese. He is 23 years old and ethnically Chinese.

3.2 Data Overview

Each consultant was given a list of twenty-six sentences to read aloud. Each sentence comprised of a carrier phrase with a target word, also known as a variable word, which varied in syllable length. A carrier phrase is a sentence used to introduce a target word. The phrase remains consistent throughout the data set while the target word is changed to test a certain variable (in this case, syllable number). Target words were chosen based on sonority, because a more sonorous word would show pitch fluctuations easier. The majority of the sentences had the target word in the middle of the utterance, so as not confuse the word boundary with the edge boundary. However, a few sentences purposefully placed the target word at the end of the sentence in order to see if there was a difference.

Sentences (1), (5), (10) and (15), are examples from the complete set of sentences found in the Appendix, used here to illustrate the concept of carrier phrases and target words. The phrase Say [insert] again is the carrier phrase for these five sentences. The target word is placed within this phrase. The target words within this set are all phonetically similar and primarily differ in syllable length. In sentence (1) the target word Min is made up of one syllable, in sentence (5), mini is made up of two syllables, in sentence (10) minimal is three syllables and, finally, sentence (15) has the four syllable word minimally. The meanings of the target words do not matter.

(1) Say Min again.
(5) Say mini again.
(10) Say minimal again.
(15) Say minimally again.
In Section 3.3 each sentence is labeled with both Ng and Chong’s Models. For Ng’s Model, each syllable is marked for tone based on its placement in the syllable, as discussed in Section 2.3. Similarly, Chong’s Model is notated using his notation of L* for pitch accents within APs, aL and Ha to bracket APs, either L- or H- to mark the ends of ips if needed and L% and H% to mark the ends of IPs.

The data in Section 3.3 is presented in a way that displays the F0 height for each syllable or prosodic unit in the sentence elicited by both consultants. The models are presented side by side for ease of comparison. The values of tone height were taken from reading spectrograms. Usually, this was done by paying attention to the pitch track generated by Praat. However, due to the background noise present in most of the recordings, this pitch track was not always a reliable indicator of tone height. In these cases, I impressionistically evaluated the sound files to identify pitch at different times. This was made easier in the stark low to high patterns the speakers regularly had in their speech.

An example of one pitch track is found in Figure 9. This is the pitch track for sentence (7) spoken by consultant RT8. The dark lines show the pitch track (F0 line) of the utterance. The first tier displays Chong’s model, the second tier displays Ng’s Model and the bottom tier has a gloss of the sentence. Chong’s Model marks the boundaries of phrases. For example, we can see the boundaries for the accentual phrase the word bordered by aL and Ha. The entire utterance is then marked by the edge boundaries for the ip and IP which are both low in this example. Ng’s Model splits the utterance into syllables and then marks each syllable with a tone. For example, the is marked with a low tone while word is marked with a high tone. In this pitch track we can clearly see that both models are able to capture the

---

8A full listing of pitch tracks for each sentence can be found in the Appendix.
repeating rising patterns on each word. This is especially apparent in the word and believe. Both words begin low and end high in the pitch track.

![Pitch Track Diagram](image)

Figure 9: Say believe again spoken by consultant RT

### 3.3 Data Description

The data in this section has been divided based on syllables. For each section we will compare Chong and Ng’s models. We will point out contradictions and any problems with how each model handles the data. However, we will not go into depth on these issues until Section 4.

#### 3.3.1 One and Two Syllable Words

For words that consist of one and two syllables, both Chong and Ng’s models are accurate, with varying levels of clarity. One thing of note, however, is that Chong explicitly accounts for function words by placing them within APs. Ng, on the other hand, does not mention them. However, in her own examples she notes them with a low tone, as I have replicated...
In order to understand how both models are able to account for one syllable words, we should examine the data in depth. Figure 10 shows the pitch track for sentence (1) Say Min again, said by consultant RT. The first tier of the figure displays Chong's Model using AM theory, the second tier displays Ng's Model, and the third tier displays a gloss of the sentence.

In Figure 10 we can clearly see the low to high pattern in both the target word Min and in the word again in the pitch track itself. This pattern is best expressed with Chong's Model which notes that the beginning of the AP is low (aL) and ends in a high (Ha). This differs from Ng's model which simply states if the syllable is high (H) or low (L). In one syllable words, such as Min, this rising pattern is not easily communicated. It is only notated as high tone, making the comparison with multisyllabic words a harder one to make.

We see this same pattern repeated in all of the one syllable word sentences, for both consultants PK and RT. The results for those sentences are reproduced below in table form. One thing to note is the phrase the word and how it behaves. In Chong's Model, it makes up one AP rather than two, as the is a function word associated with word. In Ng's Model, the representation is still accurate as we have chosen to mark function words with a low tone.

The data (1a)-(4b) list all of the sentences with one syllable target words. The data shows that both speakers produced the same intonational patterns. In addition, both models were able to predict these patterns. (1a)-(4a) apply Ng's Model's predictions to the data while (1b)-(4b) use Chong's Model's predictions. The Appendix has a listing of all pitch tracks for each sentence and consultant, annotated with both models.
Figure 10: *Say Min again* spoken by consultant RT

(1a) Ng Model  (1b) Chong Model
Say Min again.
PK H H LH
RT H H LH
Say Min again.
PK aL Ha aL Ha aL Ha H-H%
RT aH La aL Ha aL Ha H-H%

(2a) Ng Model  (2b) Chong Model
Say live again.
PK H H L H
RT H H L H
Say live again.
PK aL Ha aL Ha aL Ha H-H%
RT aL Ha aL Ha aL Ha L-L%

(3a) Ng Model  (3b) Chong Model
Say the word bee again.
PK H L H H L H
RT H L H H L H
Say the word bee again.
PK aL Ha aL Ha aL Ha aL Ha L-L%
RT aL Ha aL Ha aL Ha aL Ha L-L%

(4a) Ng Model  (4b) Chong Model
Say the word may.
PK H L H H H
RT H L H H H
Say the word may.
PK aL Ha aL Ha aL Ha L-L%
RT aL Ha aL Ha aL Ha L-L%

Similar to one syllable words, two syllable words are also well described by both models.

The previous issue of Ng’s where the single tone syllable did not accurately portray the rising
pattern of the word is no longer an issue. We can see this in Figure 11, which shows the pitch track for sentence (7) *Say the word believe again* spoken by consultant RT.

Again, the low to high pattern is easily apparent. In the target word *believe*, we start with a low tone and end in a high. It is important to note that the pitch track is poorly drawn for the last word of the sentence *again*. This is likely due to the combined problem of the loud background noise as well as the fact people tend to go creaky (which disrupts the pitch) at the ends of utterances (Ladd, 2008). This makes it difficult for Praat to track the pitch in the acoustic signal. However, when listening to the audio recordings and using impressionistic annotation, the low to high rise is easily discernible. This is still discernible even in the last word of the sentence with a possible subsequent fall due to a L- or L% boundary tone.

![Pitch Track Diagram](image)

Figure 11: *Say the word believe again* spoken by consultant RT
Sentences (5a)-(8b) include all instances in which the target word has two syllables. The patterns we have laid out previously repeat themselves. Each word has a low to high rise, with the exception of the function word the. However, this is accounted for in Chong’s Model by including it in the AP the word. In Ng’s model, we have accounted for this by marking it with a low tone.

Unlike in sentences with one syllable, Ng’s model now accurately portrays the rising action of the intonation of these sentences. Because they have two syllables, we are able to mark the initial syllable as low or mid and the final syllable high, accurately portraying the rise. We mark the syllable as a mid tone if it is stressed, such as in the word mini and living. Otherwise, we mark the first syllable as a low tone, such as in the word believe. This is another aspect in which the two models differ, as Chong’s Model predicts no difference between the intonational patterns of these words. Within the data collected it is difficult to state whether or no there is a difference in the intonation of these words.

The data (5a)-(9b) list all the sentences with two syllable target words. Again, the data shows that both speakers produced the same intonational patterns. In addition, both models were able to predict these patterns. (1a)-(4a) apply Ng’s Model’s predictions to the data while (1b)-(4b) use Chong’s Model’s predictions. One can consult the Appendix for specific pitch tracks of these sentences.

(5a)  Ng Model  (5b)  Chong Model
Say mini again. Say mini again.
PK H M H L H aL Ha aL Ha aL Ha L-L%
RT H M H L H aL Ha aL Ha aL Ha L-L%

(6a)  Ng Model  (6b)  Chong Model
Say living again. Say living again.
PK H M H L H aL Ha aL Ha aH La L-L%
RT H M H L H aL Ha aL Ha aH La L-L%
3.3.2 Three Syllable Words

In words with three syllables, issues begin to appear in Ng’s Model while Chong’s Model remains strong. Some three syllable words in CSE have a low pitch target in the middle of the word. This is allowed for in Chong’s Model with a low pitch accent (L*) but not in Ng’s. For example, in Figure 12, we see the three syllable target word believer. In Chong’s Model this is transcribed as having an initial AP low (aL) and a final AP high (Ha). However, within the word, the second syllable has a low pitch accent, shown as L*. In Ng’s model, it is impossible for a stressed syllable to have low tone. Instead, we must notate the syllable as a mid tone, regardless of the fact the pitch track clearly shows it remains a low tone. Therefore, this mid tone does not accurately represent the pronunciation of the word and is a problem in Ng’s analysis.

This low pitch accent variable occurs in some of the three syllable target words within the dataset. Examples (12b) and (13b) both show that the words believer and maybelline both have a low pitch accent (this finding is consistent between both speakers). In example (11b), however, we see that consultant PK has a low pitch accent in the middle syllable of livable but RT does not. Examples (10b) and (14b) show two three syllable words, minimal and
Figure 12: *Say the word believer again* spoken by consultant PK

Imagine which have no low pitch accent for both speakers. For these latter two sentences, Ng’s Model works equally as well as Chong’s Model. There is a debate surrounding the existence of the low pitch accent ($L^*$). Chong states that it is optional and that it seems to only occur in the prominent word of the sentence, although further research is needed to fully understand it (Chong, 2012).

The rest of the data within these sentences remains consistent with two syllable and one syllable sentences. The words say and again both have an initial low and final high. The phrase *the word* continues to work as an AP$^9$.

$$
\begin{array}{cccc}
(10a) & \text{Ng Model} & (10b) & \text{Chong Model} \\
\text{Say} & \text{minimal} & \text{again.} & \text{Say} & \text{minimal} & \text{again.} \\
PK & H & M M H & L H & aL & Ha & aL & Ha & aL & Ha & L-L% \\
RT & H & M M H & L H & aL & Ha & aL & Ha & aL & Ha & L-L% \\
\end{array}
$$

$^9$In sentence (11), RT’s data is difficult to discern. Because of this, the missing values are marked with an $X$. 

25
3.3.3 Four or More Syllable Words

There are only two instances in which the low pitch accent is present in a four syllable word. In example (15b), consultant PK pronounces minimally with a low pitch accent on the third syllable but RT does not. Conversely, in sentence (17b) consultant RT pronounces imagining with a low pitch accent on the third syllable while PK does not. In all other sentences with four syllable target words, (16b)’s believable, (18b)’s malaria and (19b)’s Romania, neither speaker pronounces the word with a low pitch accent.

(15a) Ng Model (15b) Chong Model
Say minimally again. Say minimally again.
PK H L M M H L H aL Ha aL Ha aL Ha L-L%
RT H L M M H L H aL Ha aL Ha aL Ha L-L%

(16a) Ng Model (16b) Chong Model
Say the word believable again. Say the word believable again.
PK H L H L M M H L H aL Ha aL Ha aL Ha L-L%
RT H L H L M M H L H aL Ha aL Ha aL Ha L-L%

(17a) Ng Model (17b) Chong Model
Say the word imagining again. Say the word imagining again.
PK H L H L M M H L H aL Ha aL Ha aL Ha L-L%
RT H L H L M M H L H aL Ha aL Ha aL Ha L-L%
In multisyllabic words greater than four, more problems begin to emerge for both Chong’s and Ng’s Models. In Figure 13 we see the five syllable target word *livability* spoken by consultant RT. Here there is a low pitch accent found on the second syllable. This is not predicted accurately in either model. In Chong’s model, the L* can only be placed on a stressed syllable, and the second syllable of *livability* is unstressed. In Ng’s model, a low pitch accent within a word is not possible. In both cases, this gives the false representation that the entire word is a continuous rising pattern. On the other hand, other five syllable words behave as would be predicted by both models, such as the word *radiology* in (25).

This same pattern seems to repeat in other six syllable words. For example, in Figure 14 we see consultant PK’s pronunciation of *memorabilia*. Ng’s Model does not accurately predict the sudden drop in tone in the middle of the word. Chong’s Model could make use of the low pitch accent, however, it does not accurately predict the pronunciation. The first three syllables of the word have an initial low and then rise to a high. The fourth syllable has a low pitch target and the rest of the word displays the rising action once more. This initial rise during memora-, before the L*, is not accurately portrayed in Chong’s Model.

In addition, the majority of the sentences with six syllable words contain this pattern. For the words, *livability* (20), *unimaginable* (21), *developmentally* (23), *memorabilia* (24) and *biodegradable* (26), both speakers have an initial rise in the word, followed by a low
pitch target within the same word, which seems to reset the low to high rise. In addition, consultant PK has this in the word developmentally in example (22).

Because of this consistent issue in both models for multisyllabic words greater than four syllables, it is necessary to revise one of the models in order to account for this. This will be discussed in Section 4. In addition, it is difficult to show the errors in the prediction in table format. As such, refer to the appendix to see the pitch tracks for sentences (20)-(26).

4 Analysis

It has become apparent by looking at the data, that both Chong and Ng’s models are not entirely adequate for describing CSE’s intonational system. For one, Ng’s model only works for two syllable words and greater if the word does not contain a low pitch accent. Because it is problematic for the majority of the data with these longer words, I will be setting it
aside. Chong’s model on the other hand, is able to describe almost all of the data accurately, including the examples where Ng’s Model fails. However, Chong’s model contains gaps in respect to multisyllabic words greater than four syllables. Because of this, I propose to modify Chong’s existing model to fill in these gaps.

4.1 Prefixes as Unique Prosodic Units

There are two modifications Chong’s Model needs in order to accurately represent the data. The first is that prefixes in CSE are treated as their own AP phrase. Others have discussed how, in many languages, prefixes “fail to incorporate into the prosodic word to which they attach” (Peperkamp, 1999). In other words, that prefixes form their own unique prosodic unit from the stem that hosts them, morphologically. In the case of CSE, I argue that prefixes are their own accentual phrase. This simple addition would account for words like *un-imaginable* and *bio-degradable*. 
In Figure 15 there is a clear rise on the syllable un followed by an immediate fall as the rest of the word begins. Chong’s L* would not be able to account for the rising pattern on
the prefix. It is not just a low pitch accent but the restarting of a new accentual phrase. Furthermore, there seems to be a perceptible juncture between the two syllables in question. We see a similar process in Figure 16 spoken by consultant PK. Here, there is a clear rising tone on the prefix bio with a sudden dramatic fall on the next syllable. It is more salient to mark this change by using two separate APs rather than a low pitch accent. Thus, I appeal to the idea that prefixes are prosodically separate, and propose that prefixes in the intonational system of CSE should be treated as their own accentual phrase. This modification is compatible with Chong’s Model, as it only involves changing the definition of what constitutes an accentual phrase.

4.2 Multisyllabic Words and the Accentual Phrase

In addition to the tendency for prefixes to be unique prosodic words, we must propose another modification in order to account for words greater than four syllables. I propose that these longer multisyllabic words which behave differently actually contain two accentual phrases, as the intonation pattern of low to high seems to repeat itself. Whether or not these are indeed APs and not some other smaller prosodic unit is unknown. However, it does seem to be necessary to break down these multisyllabic words further in some way in order to account for the pronunciation we are hearing, both in terms of intonation and juncture. In both six syllable and five syllable words, the AP break appears in the middle of the word. Again, this would not change Chong’s model drastically, but rather modify how APs are defined within it.

For example, we see this in the six syllable words memorabilia and developmentally. In Figure 17 we see a rise in the first half of the word, followed by a break in the speaker as
the AP seems to restart the intonation pattern. This is heard in Figure 18 as well. When listening to the recording the repeating rising pattern is clearly apparent. In both cases, there seems to be a slight perceivable break in between the proposed APs in the pronunciation, giving further evidence for the existence of a boundary within words.

Figure 17: Say memorabilia again spoken by consultant PK
In the five syllable word *livability* we have a similar problem with the prediction. Here, it seems the second syllable of the word remains relatively low. We cannot use a L* because it is not the stressed syllable of the word. When listening to the recording we can hear a low to high pattern occurring on the initial two syllables, and again on the latter three. In this case, the first AP surrounds the first two syllables of the word *li*-va. The last three syllables make up the next AP *bi-li-ty*. We can see this in Figure 19. It is important to note this is the only example of this pattern in five syllable words within the available dataset.
4.3 Ternarity

Within the data set there seems to be a preference towards *ternarity* in accentual phrases of CSE. A ternary system is one in which prosodic units are grouped in threes, in contrast to a binary system which groups prosodic units in two. Researchers have found more evidence for binarity in prosodic structure than ternarity. Most syllable systems use binary feet, composed of two syllables. (Goedemans, 2013). The fact that CSE may have a preference for ternarity in building APs out of syllables requires further examination and discussion. CSE’s AP length preference is recorded in Table 3. The first column lists the number of syllables in the target word. The corresponding column shows possible AP configurations, separated by backslashes ( / )\(^{10}\).

\(^{10}\)Note that these are for attested words without prefixes.
There is a tendency in multisyllabic words greater than four syllables to break into two APs, as discussed in Section 4.2. Within this tendency a pattern preferring three syllable units becomes apparent. In six syllable words we see cases in which the word splits into two three syllable accentual phrases. In five syllable words we see a similar split occurring, manifesting itself as one two syllable AP followed by a three syllable one. Target words with less syllables are never split into multiple accentual phrases. In a binary system, one would expect a different outcome. In six syllable words, we would see a preference toward three two syllable APs. In five syllable words, we would have two two syllable APs and a lone single syllable AP. Finally, we would also see two two syllable APs in four syllable words. In CSE, four syllable words never seem to have multiple APs within them.

Not only do we see a preference for ternarity in this data, but also a preference for right-alignment. Five syllable target words within the data set that have two APs always have the three syllable AP to the right of the word. It would be interesting to investigate if and how right alignment manifests itself in other parts of Colloquial Singaporean English. The data set within this paper is very limited and only has two consultants. It is important to investigate this matter further and see if it is a consistent pattern in more speakers alongside more data. I will not be looking into it further other than to comment on it's tentative existence here.

Another key fact from Table 3 is found in target words of four syllable length. Within the data collected, these multisyllabic words are never split into two APs. One would predict the possibility of a one syllable AP followed by a three syllable unit for these words. However, this seems to be impossible. This may hint that CSE disfavors one syllable AP, and only uses them when it is necessary—for one syllable words and prefixes. This could also lead into
an investigation on the ordering for these AP rules. An examination of this data through
the lens of Optimality Theory would be an interesting topic for a future paper.

<table>
<thead>
<tr>
<th># of Syllables in Target Word</th>
<th># of Syllables in AP(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>(σσσ)(σσσ)/(σσσσσσσ)</td>
</tr>
<tr>
<td>5</td>
<td>(σσ)(σσσ)/(σσσσσσσ)</td>
</tr>
<tr>
<td>4</td>
<td>(σσσσσ)/*(σ)(σσσ)</td>
</tr>
<tr>
<td>3</td>
<td>(σσσ)</td>
</tr>
<tr>
<td>2</td>
<td>(σσ)</td>
</tr>
<tr>
<td>1</td>
<td>(σ)</td>
</tr>
</tbody>
</table>

Table 3: Target Words and Corresponding Syllable AP Length

4.4 Exceptions

There are, however, exceptions to these modifications. For example, there are certain cases
in which the target word developmentally does not exhibit this two AP pattern. In Figure
20 we see consultant RT’s pronunciation of the target word does not have two APs, nor
does it have a low pitch accent. Instead, it has a consistent rising action throughout the
entire word. One possible explanation is rate of speech. It could also be due to the fact the
speaker seemed a bit distracted as he pronounced the data, hinted at with the unusual pause
between the target word and the word again. This could also be due to speaker variation.
Lastly, it could just be an exception to the rule. However, it is important to acknowledge it.

As it stands now, these two modifications surrounding prefixes and multisyllabic words
help strengthen Chong’s current model and better represent the intonation system of Col-
loquial Singaporean English as we know it. His model remains relatively in tact, with only
a call for further investigation on what exactly constitutes an accentual phase. These modi-
fications are in need of more data to be fully substantive. Moreover, more data is necessary
to explore CSE’s apparent tendency for ternary systems.

5 Conclusion

In Colloquial Singaporean English there is a general intonational rise that occurs over the course of a single prosodic unit, usually the size of a content word. This is found with interspeaker consistency. This finding confirms both Chong and Ng’s Model, as each predict this rising pattern. One gap in the literature, however, is how multisyllabic words behave within this system. In this thesis, I compared both models and their ability to predict longer multisyllabic words accurately in order to improve upon this subject.

From the data collected it is apparent that Chong’s Model is currently the most accurate at representing the dataset. Ng’s Model is inadequate when describing single syllable words
and multisyllabic words with a low pitch accent on the stressed syllable. Chong’s Model remained consistently able to predict multisyllabic words of longer length, but is in need of minor modification in relation to what an accentual phrase is consistent of.

In Section 4, I proposed two modifications to Chong’s Model in order to better describe the data set. This included a modification that treated prefixes as their own prosodic unit separate from the word they are attached too. In addition, I proposed that larger words, such as six syllable words, may actually contain two APs rather than one. However, in order to substantiate these claims more data is necessary.

One avenue to take in investigating the nature of APs in CSE is to look at boundaries within words such as cannot. Cannot is treated as its own word in Ng’s analysis, as stated in Section 2.3. It would be interesting to look into this further and see if other similar words behaved in this way, such as will not. This may shed more light into the nature of AP boundaries.

An important aspect of this research to note that has not been brought up is the behavior of the word say. There are times in which it is in the same AP as the following word, as there is no perceptible rise on it, pointed out to me by Chong. This is problematic, as it is its own content word and should be have consistently throughout the data. This behavior could be explained by the monotonous task of the elicitation, as each sentence begins with the word say every time. It could also hint again towards CSE’s dispreference toward single syllable APs. It is an aspect of this research that deserves further probing.

During this investigation, a tendency towards ternarity made itself apparent in CSE’s intonation system. In multisyllabic words with two APs, the split always occurred so that the final three syllables remained a unit. This preference for three syllable units seems
typologically unusual and deserves its own further exploration. Does this preference exist in other instances of CSE? If so, where? Is this another borrowing from a language it is in contact with, or did it come about by itself? In addition, there seems to be a preference toward right alignment as well. Is this apparent in other aspects of the language? These questions are outside the scope of this paper but would make excellent topics for further investigation.

Furthermore, another important question is still left unanswered. What type of language is Colloquial Singaporean English? It does not behave as a lexical tone language like Mandarin Chinese, as it does not have lexical minimal pairs differentiated only by tone. In addition, although stress is important to it’s intonation system (Ng uses it to ascribe tone to a syllable, Chong uses it in his definition for low pitch accents), it does not seem to have any prominence lending abilities like stress-accent languages have. This leaves CSE as a pitch-accent language, similar to Japanese. The low pitch accent in Chong’s model does seem to hint at this, as it is ascribed to the lexical word based on its stress pattern. However, there is too much variability within the speakers to see this as a consistent aspect of the language. This leads me to believe that Colloquial Singaporean English exists outside of these three options. This could be because it has been influenced by so many languages of different intonational type. More data is needed to investigate this matter further.

It is important to revisit the conclusions of this paper with more data, specifically data of longer multisyllabic words. In my initial investigation, I elicited a large array of target words with varying syllable length. In the future I would look more closely at just longer multisyllabic words to search for more patterns and see if there is a bias towards ternarity. In addition, it would be important to consult with more speakers. It would also be interesting to
look at other kinds of morpheme boundaries and how they play into the intonation systems. Finally, another avenue of thought would be to examine an OT analysis of accentual phrase boundary ranking. Especially to see how much one syllable APs are dispreference. One could test this with short words with prefixes and so forth.

In searching for answers to multisyllabic words in CSE’s intonation system we have discovered an array of more questions, from ternarity to AP boundaries. All of these questions are deserving of answers, but are unfortunately outside the scope of this paper. More data is necessary to continue the search for these answers.
Sentence (1) spoken by PK

Sentence (1) spoken by RT
Sentence (2) spoken by PK

Sentence (2) spoken by RT

42
Sentence (3) spoken by PK

Sentence (3) spoken by RT
Sentence (4) spoken by PK

Sentence (4) spoken by RT
Sentence (5) spoken by PK

Sentence (5) spoken by RT
Sentence (6) spoken by PK

Sentence (6) spoken by RT
Sentence (7) spoken by PK

Sentence (7) spoken by RT
Sentence (8) spoken by PK

Sentence (8) spoken by RT
Sentence (9) spoken by PK

Sentence (9) spoken by RT
Sentence (10) spoken by PK

Sentence (10) spoken by RT
Sentence (11) spoken by PK

Sentence (11) spoken by RT

51
Sentence (12) spoken by PK

Sentence (12) spoken by RT
Sentence (13) spoken by PK

Sentence (13) spoken by RT
Sentence (15) spoken by PK

Sentence (15) spoken by RT
Sentence (18) spoken by PK

Sentence (18) spoken by RT
Sentence (19) spoken by PK

Sentence (19) spoken by RT
Sentence (20) spoken by PK

Sentence (20) spoken by RT
Sentence (21) spoken by PK

Sentence (21) spoken by RT
Sentence (22) spoken by PK

Sentence (22) spoken by RT
Sentence (23) spoken by PK

Sentence (23) spoken by RT
Sentence (25) spoken by PK

Sentence (25) spoken by RT
Sentence (26) spoken by PK

Sentence (26) spoken by RT
REFERENCES


Chong, Adam & James Sneed German. 2015. Prosodic phrasing and F0 in Singapore English.


