References

Abbott, Gerry. 1979. Intelligibility and Acceptability in Spoken and Written Communication. *ELT Journal* 33. 3:168–175.

Brown, Adam. 1988. A Singaporean Corpus of Misspellings: Analysis and Implications. *Journal of the Simplified Spelling Society* (UK) 3. 4–10.

Ho Mian Lian. 1997. This is Chickenfits: The Effect of Pronunciation on Writing. *RELC Journal* 28. 2:22–34.

Randall, Mick. 1997. Orthographic Knowledge, Phonological Awareness and the Teaching of English: An Analysis of Word Dictation Errors in English of Malaysian Secondary School Pupils. *RELC Journal* 28. 2:1–21.

Rickard Liow, Susan. 1998. Phonological Awareness in Multilingual Chinese Children. *Applied Psycholinguistics* 19. 3:339–362.

Siew Sook Yee. 1984. An Investigation of the Clues Given by Misspellings for the Phonemic System of the Singaporean Pronunciation of English. Unpublished Honours Dissertation, Department of English Language & Literature, National University of Singapore.

Tesdell, Lee S. 1987. Patterns in the Spelling Errors of Non-native Speakers of English. *Guidelines* (RELC) 9. 1:80–84.

Yeo, Josephine. 2001. Potential Influences of 'Mother-tongue' on the Written English of Secondary School Students in Singapore. Unpublished Honours Dissertation, English Language & Literature Academic Group, National Institute of Education, Singapore.

Yeo, Josephine & Deterding, David. 2003. Influences of Chinese and Malay on the Written English of Secondary Students in Singapore. In *English in Singapore: Research on Grammar*, 2003. Edited by David Deterding, Low Ee Ling & Adam Brown. Singapore: McGraw-Hill. pp. 77–84.

Varieties of English in Southeast Asia and Beyond
Azirah Hashim and Norizah Hassan (eds.)
Kuala Lumpur: University of Malaya Press
2006

9

IS THE STRESSED SYLLABLE STRESSED?! THE PERCEPTION OF PROMINENCE IN SINGAPORE ENGLISH

Tan Ying Ying

Standard Southern British English has always been the yardstick used in Singapore English (SE) research, and this is especially true in the research of the prosody of SE. Tongue (1979), Tay (1982) and Platt and Weber (1980), among many other works, are representative of this comparative approach to the study of SE prosody. This approach compares SE and British English (BrE), listing the differences between these two varieties. Such a method of analysis is what Mohanan (1992:111) labels as the "parasitic approach" – an approach that describes features in SE as 'errors' or 'deviations' from the standard norm. This approach implicitly reinforces the view that SE is an imperfect or imprecise copy of the 'original'. Kachru (1979:7) refers to this penchant for using an "unrealistic reference to a model" in the study of a new variety of English as a "sin of exhibiting language colonialism".

Many recent researchers (e.g. Deterding, 1994b; Low, 1994, 1998; Low and Grabe, 1999) embrace the idea that SE is an autonomous language. Their methodology of research, however, cannot escape a comparison between SE and BrE. This makes them look very much like "those who overtly reject colonialism and assert the independent status of non-native system [but] are implicitly colonialist when it comes to linguistic descriptions of the structure of non-native varieties" (Mohanan, 1992:113). In the description of the prosody of SE, few linguists have studied SE in its own right, without comparing and referring SE to the British variety. This article seeks to break away from this comparative tradition in its investigation of prominence[1] in SE and aims to look at this phenomenon from its own linguistic patternings and

observations, without a presupposition of theories or assumptions that would allude to the British standard of understanding prosody.

Most of the existing studies on SE stress[2] focus on word-stress placement. Most of these studies work on comparing word stress patterns in SE to those in BrE (Tongue, 1979; Platt and Weber, 1980; Tay, 1982; Alsagoff, 1984; Ng, 1985; Chua, 1989; Sng, 1991; Deterding, 1994a; Deterding and Hvitfeldt, 1994; Bao, 1998; Low, 1998; Low and Grabe, 1999). In these past studies, the findings on the stress patterns in SE are based on the researchers' own perceptions of prominent syllables. There is a major assumption that these researchers are making, i.e. that the perception of prominence in SE is the same as the perception of prominence in BrE, and that BrE speakers would perceive prominence in SE as how they would in BrE. Tongue (1979) and Platt and Weber (1980), for example, in their experiments, use their own judgements to determine the stressed syllables in their sample. It is important to note that researchers like Tongue and Platt and Weber are BrE speakers and therefore, the judgements of SE prominence are based on BrE perceptions. As Tay (1982) points out, British listeners might perceive prominence differently from SE listeners. A higher pitched syllable, while it might sound prominent to the BrE listener, might not be prominent for a SE listener. It is thus inappropriate to assume that SE speakers have the same perceptual cues for prominence as those of BrE speakers. Before an accurate analysis of SE stress patterns can be carried out therefore, there is a need to go back to the basics, returning to question the original point: what *is* prominence in SE?

For Fry (1955, 1958, 1965), the perception of prominence denotes a complex of perceptual physical dimensions. He believes that there are four physical factors that are important in influencing one's judgement of prominence. The listener perceives prominence objectively by relying on the factors, namely, (1) the length of syllables, (2) the loudness of the syllables, (3) the pitch of the syllables and (4) the vowel qualities of the syllables. Fry's (1955, 1958, 1965) series of experiments sparked off a string of other similar works on different languages, contributing tremendously to the understanding of the perceptual nature of prominence. Putting the experiments together, he concludes that fundamental frequency is the most dominant perceptual cue, followed by intensity, then duration.

Other similar studies, for example, Bolinger (1958), using both natural and artificial speech, also come to the same conclusion that the primary cue to prominence is pitch. While he regards duration as an equally important perceptual cue, he rejects the notion that amplitude has a crucial role to play as an effective perceptual cue. Morton and Jassem (1965), using nonsense words, /sisi/, /sɔsɔ/ and /sasa/ as test items, note that a raised F_0 is more effective as a perceptual cue than a lowered one, and that the more intense and longer a syllable is, the more likely it is to be marked as prominent. The most important finding is that variations in F_0 produce far greater effects in the listeners' judgements than duration or amplitude, seemingly showing the importance of pitch as the dominant perceptual cue for prominence.

Research on other languages also find F_0 to be the overriding perceptual cue for prominence. Jassem (1959), Jassem *et al.* (1968) and Awedyk (1986) conclude from their experiments that Polish listeners take F_0 to be the dominant perceptual cue for prominence. Janota (1979) notes that in Czech, changes in F_0 are predominant factors in the listeners' perceptions of prominent syllables. Westin *et al.* (1966) study on Southern Swedish also finds F_0 to be the main perceptual cue for prominence. For the Estonians, the findings are the same, as reported by Eek (1987). For the Russians however, as Eek (1987) reports, duration, not F_0, serves as the leading parameter. In Japanese, duration plays no role in stress production or perception because a quantity opposition in vowel length is phonemic (Mitsuya and Sugito, 1978; Beckman, 1986). Amplitude, believed to be an important acoustic cue, is later shown to have little influence in stress perception in Japanese (Weitzman, 1969; Beckman, 1986; Beckman and Pierrehumbert, 1986; Hasegawa and Hata, 1992). It seems, therefore, that in Japanese, F_0 is also the overriding perceptual cue for stress.

The Experiment

In the investigation of prominence in SE, a perception test is designed and conducted. The experiment is in two main parts. The experiment is intended to show to what extent each of the three parameters (fundamental frequency, duration and amplitude) is, or may be responsible for the impressions identified by these SE subjects as *prominence*. For the first part, the main purpose is to determine if the Singaporean subjects use higher or lower pitch, greater or less intensity and shorter or longer vowel duration to determine stress. Having established that, the second part of the experiment concentrates on determining the relative strengths of the perceptual cues. In other words, when faced with a choice between two syllables, one of which is longer and the other, louder, for example, which syllable the subjects would choose as the more prominent one.

The materials consist of three utterances. They are:

(1) I **see sea** creatures. [si:si:]
(2) I **saw saw** blades. [sɔsɔ]
(3) He'll **sue Sue** later. [su:su:]

The same vowel is used in both syllables to ensure that there are no other phonetic considerations that can influence the subjects' choice besides the acoustic correlates themselves. The syllables also have the same structure, [CV], with the same consonant, [s] used in all the syllables to avoid the possibility of different consonants affecting the intrinsic phonetic properties present in the following vowel. Three different, "extreme" vowels are used.

The choice of [ɔ:] was to replace [a:], which could not be used in this case, as [sa:] is not a word in English. The syllables/words tested are sounds and sound sequences that can be found in English to make sure that the test stimuli are as close to real speech situation as possible.

Synthesis

The three test sentences in the test material were 'spoken' by a computer-generated speech synthesis program. This original stimulus was taken from *The Festival Speech Synthesis System: University of Edinburgh*[3]. For the purpose of this research, the 'speaker' that speaks Standard Southern British English is chosen as the provider of the utterances. The advantage of using this system is that the speech is machine-generated, so it is much easier to control the utterances, as compared to getting humans to read the utterance.

The 'speaker' was made to say each word separately, so as to avoid any intonation or rhythmic pattern interfering with the test stimuli. The words were later put together using Praat[4] Version 3.0, with each word being put at an equal length to the next. This is to make sure that no unnecessary pauses or breaks would interfere with the perception of these sentences.

The test syllables were synthesised on Praat, with a predetermined set of individual parameters. The vowel, not the whole syllable was synthesised. Though the syllable is the basic unit of perception (Ladefoged, 1967), it is the vowel that takes the bulk of the suprasegmental load (Studdert-Kennedy, 1976:270) as the vowel is "relatively stable, high in energy, and spectrally compact", and it allows the speaker to display variations in fundamental frequency, duration and intensity to "offer possible contrasts in stress and intonation" (*ibid.*).

The basic values for F_0, amplitude and duration chosen for each vowel are 120Hz, 70dB and 0.2sec respectively, as they are the closest to natural speech (of an average man), spoken in a relatively quiet environment, and at a relatively normal speed[5]. The four words in each test sentence have the following values, and the sentence, as seen below is the **base form** template for each sentence:

Word-1	Target word	Target word	Word-4
↓	↓	↓	↓
F_0 = 98Hz	F_0 = 98Hz	F_0 = 120Hz	F_0 = 120Hz
Amplitude = 60dB Vowel length = 0.08sec	Amplitude = 70dB Vowel length = 0.2sec	Amplitude = 70Db Vowel length = 0.2sec	Amplitude = 60dB Vowel length = 0.08sec

As can be observed, the vowels in the two words wrapping the test words, i.e. the first and the last word in the sentence have lower values in F_0, vowel length and amplitude, compared to that of the target words, so as to make sure that the subjects would concentrate solely on the target words.

The test words in the base form were then manipulated in four levels[6] – for all three parameters of F_0, amplitude and duration. The set of values were derived on the basis that when two test items were compared, they were audibly distinguishable[7]. The values chosen for the step manipulations are:

F_0	: 100Hz, 110Hz, 130Hz and 140Hz.
Amplitude	: 60dB, 65dB, 75dB and 80dB.
Vowel length	: 0.10sec, 0.15sec, 0.25sec and 0.30sec.

For the first part of the analysis, consisting of 36 utterances (12 utterances per vowel set), the main aim is to determine if the Singaporean subjects use higher or lower pitch, greater or less intensity and shorter or longer vowel duration to determine prominence. The first test syllable was kept at the base form (120Hz, 70dB, 0.2sec), and the second test syllable had one parameter being manipulated at one time, with the other two parameters being held constant. The following is an illustration of the manipulation of duration, as an illustration[8]:

Word-1	Target word	Target word (manipulated)	Word-4
↓	↓	↓	↓
F_0 = 98Hz	F_0 = 120Hz	F_0 = 120Hz	F_0 = 98Hz
Amplitude = 60dB Vowel length 0.08sec	Amplitude = 70dB Vowel length = 0.2sec	Amplitude = 70dB Vowel length = 0.3sec	Amplitude = 60dB Vowel length = 0.08sec

The comparison therefore is now between a shorter (first target word) and a longer (second target word).

In the investigation of the relative strengths of the parameters, 36 utterances, similarly 12 utterances per vowel set were synthesised, this time, having two parameters manipulated at the same time, with both the first and the second vowel each having one parameter manipulated, and the remaining parameter kept constant[9]. For example, when one wants to compare the relative strengths of amplitude and F_0, the sentence and its component words will have the following values:

Word-1	Target word (manipulated)	Target word (manipulated)	Word-4
F_0 = 98Hz	F_0 = 140Hz	F_0 = 120Hz	F_0 = 98Hz
Amplitude =60dB Vowel length = 0.08sec	Amplitude =70dB Vowel length = 0.2sec	Amplitude = 80dB Vowel length = 0.3sec	Amplitude = 60dB Vowel length = 0.08sec

The comparison is now between a higher but softer (first target word) and a lower but louder (second target word).

All 72 utterances were randomised, and together with eight filler utterances placed at the beginning, middle and end, the set of 80 utterances was recorded into a cassette tape.

150 undergraduates[10] with normal hearing from the National University of Singapore, aged between 19-27, partook in this experiment. The subjects were given a questionnaire which asked about the subjects' linguistic profile[11] before they were played the perception test. This was to determine that the subjects were native Singaporean speakers, and were bilingual speakers of English and their respective Mother Tongue[12].

The subjects listened to the tapes in groups or individually. Each listening session had not more than six people at one time. The test was held in the sound-proof Phonetics Laboratory in the National University of Singapore. The tape was played to them using a good quality tape recorder. They were given instructions in the questionnaire to listen to the tape carefully, and to tick the word in the sentence they felt was *prominent*. They were also given the choice to leave the option blank if they could not decide which word was the more prominent one. Each sentence was played twice. The whole listening test lasted 10 minutes.

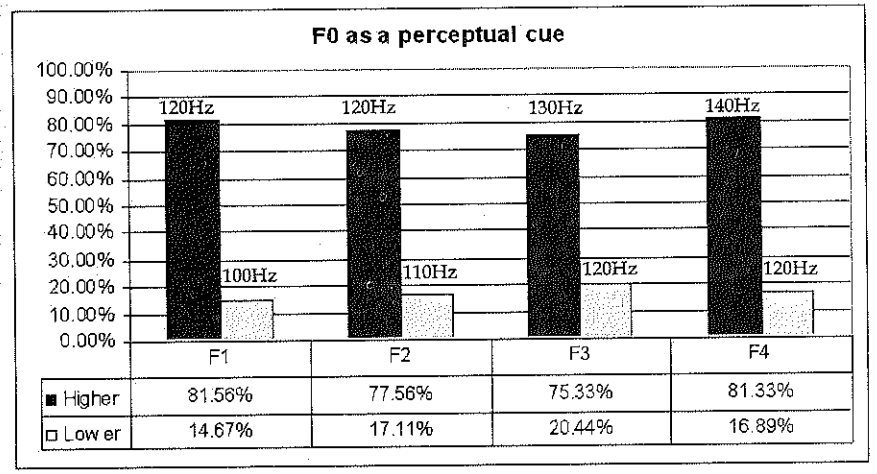
The main concern of the analysis is to establish the percentage of listeners in their judgements of prominent syllables. The following section will discuss F_0, amplitude, and duration as perceptual cues.

Results

F_0 as a Perceptual Cue

The subjects' choice of the more prominent vowel, in this section, is dependent upon a difference in fundamental frequency. Amplitude and duration are kept at the same value of 70dB and 0.2sec respectively for both vowels in each set of test words, in all four steps of manipulation. Figure 1 shows the responses of the subjects when faced with a choice between a higher pitched and lower pitched syllable.

Figure 1: Responses of subjects when choosing between higher pitched and lower pitched syllables

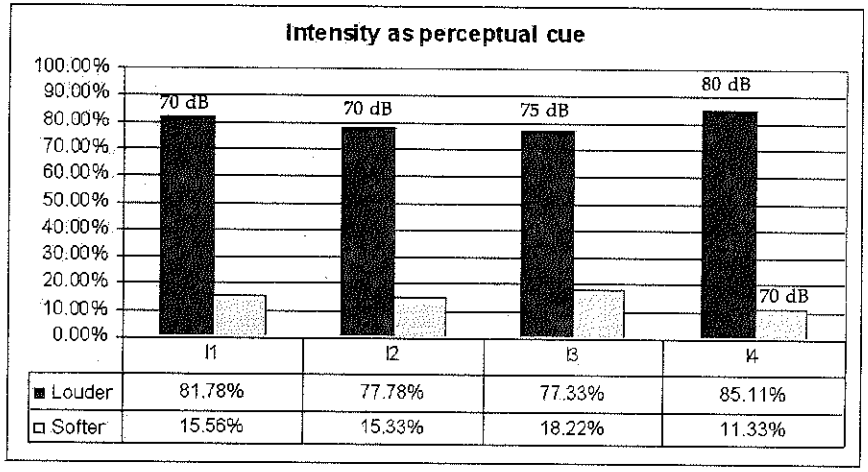


As can be seen from Figure 1, for all four levels of manipulations, more than 75% of the subjects choose the higher pitched syllable as the more prominent one. When the difference in F_0 between the two syllables is 20Hz (as in *F1* and *F4*), the percentage of subjects choosing the higher syllable goes over 80%, showing the extent to which a pitch differential is essential in allowing for a keener perception of prominence.

Intensity as a Perceptual Cue

The subjects' choice of the more prominent vowel, in this section, is solely dependent upon a difference in amplitude. F_0 and length are kept at the same value of 120Hz and 0.2sec respectively for both vowels in each set of test words, for all four steps of manipulation. Figure 2 presents the judgements of the subjects when faced with a choice between a louder and softer syllable.

Figure 2: Responses of subjects when choosing between louder and softer syllables

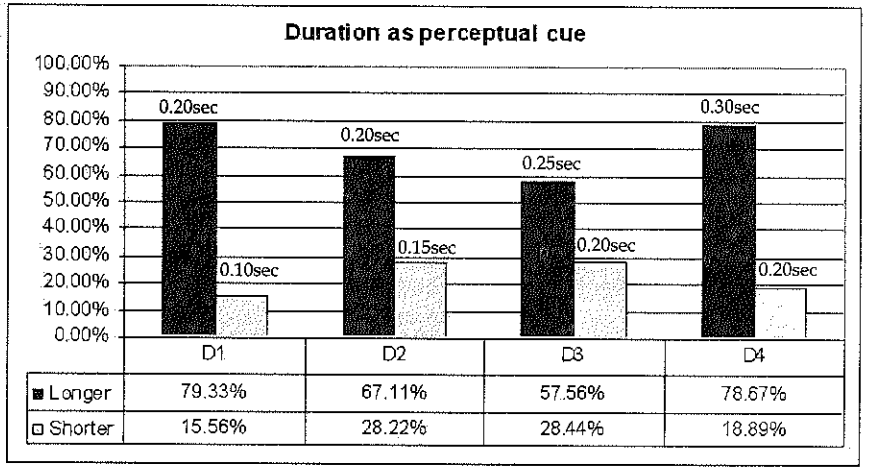


As can be observed from Figure 2, it is consistent across all four levels of manipulation that the test word containing the louder vowel is perceived as the more prominent item, with more than 75% of the subjects in each group choosing the louder syllable as the prominent syllable and this is true across the four different manipulations of amplitude.

Duration as a Perceptual Cue

Amplitude and F_0, in these 4 manipulations, are kept at the same value of 70dB and 120Hz respectively for both vowels in each set of test words, and thus the subjects' choice of the more prominent vowel is solely dependant upon a difference in vowel length. Figure 3 shows the results of the perception test when the subjects are presented with a target word containing a longer vowel compared to one with a shorter vowel.

Figure 3: Responses of subjects when choosing between longer and shorter syllables



As can be observed from Figure 3, when the difference between the two vowels is 0.10sec (in the case of *D1* and *D4*), close to 80% of the subjects agree that the test word containing the longer vowel is the more prominent item. However, when the difference between the two vowels is only 0.05sec, as can be seen in *D2* and *D3*, the decision becomes less obvious, with less than 70% of the subjects choosing the word containing the longer vowel as the more prominent one, and in the case of *D3*, only a mere 57% of the subjects choosing the longer syllable as the more prominent syllable.

While there is a general consistency in the subjects' choice for the longer syllable as the more prominent syllable, in comparison to what is seen in pitch and intensity as a perceptual cue for prominence, what is seen in duration as a perceptual cue is less clear-cut, with seemingly more room for diverse perceptions.

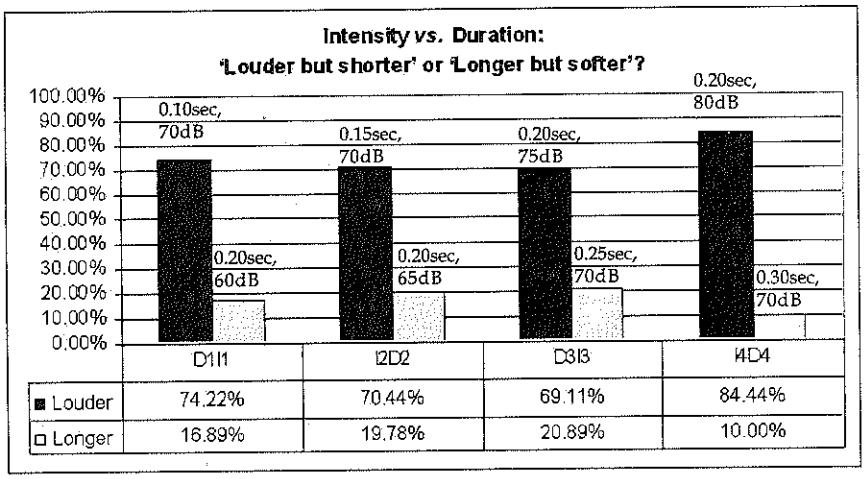
Comparison between the Relative Strengths of Each Perceptual Cue

As mentioned earlier in the paper, there are two parts to the experiment. From the last section, it has been successfully established that the subjects tend to perceive a higher pitched, louder and longer syllable as the prominent one, compared to their lower pitched, softer and shorter counterparts. In this section, comparisons between the relative strengths of each perceptual cue, F_0, amplitude and duration will be made.

Intensity vs. Duration: 'Louder but Shorter' or 'Longer but Softer'?

In this section, comparison is made between a test word that is longer and softer, versus one that is shorter but louder. For the test word in which vowel length is manipulated, amplitude and F_0 is kept constant at 70dB and 120Hz. In the test word in which the amplitude is being manipulated, F_0 and duration is kept constant at 120Hz and 0.20sec. Figure 4 shows the results of the perception test when the subjects are presented with a louder but shorter target word and a longer but softer target word.

Figure 4: Responses of subjects when choosing between louder but shorter and longer but softer syllables



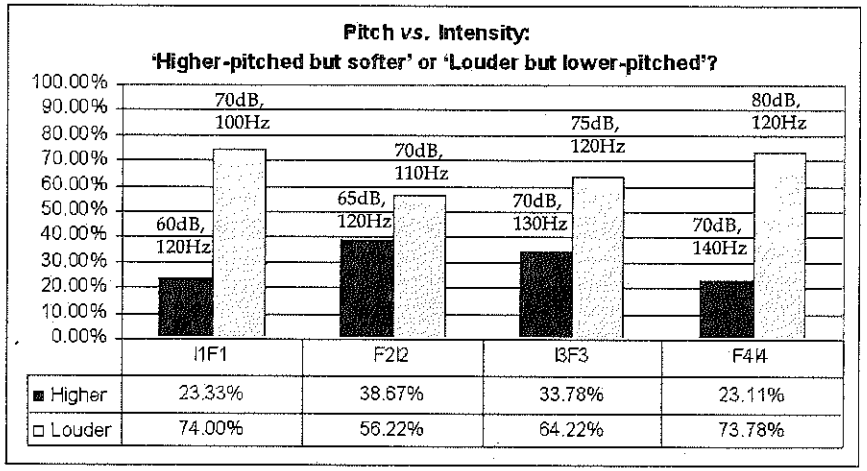
As can be seen from Figure 4, when the subjects are to choose between a louder but shorter test word and a longer but softer test word, their choice is for the former, with up to 85% of the subjects choosing the louder target word over the longer one, as in the case of *I4D4*, though the percentage sees a dip to about 70% in *D3I3*, when the difference in amplitude between the two target words is only 5dB. On the whole however, intensity is more dominant than length for the SE subjects, for all four levels of manipulations.

Pitch vs. Intensity: 'Higher Pitched but Softer' or 'Louder but Lower Pitched'?

In the test word in which F_0 is manipulated, amplitude and duration is kept constant at 70dB and 0.20sec. In the test word in which the amplitude is being manipulated, F_0 and duration is kept constant at 120Hz and 0.20sec.

The comparison, in this section, is between the word that is higher pitched but softer, versus one that has a lower pitch but louder – a choice between intensity and pitch as perceptual cues. Figure 5 presents the results of the perception test.

Figure 5: Responses of subjects when choosing between higher pitched but softer and louder but lower pitched syllables

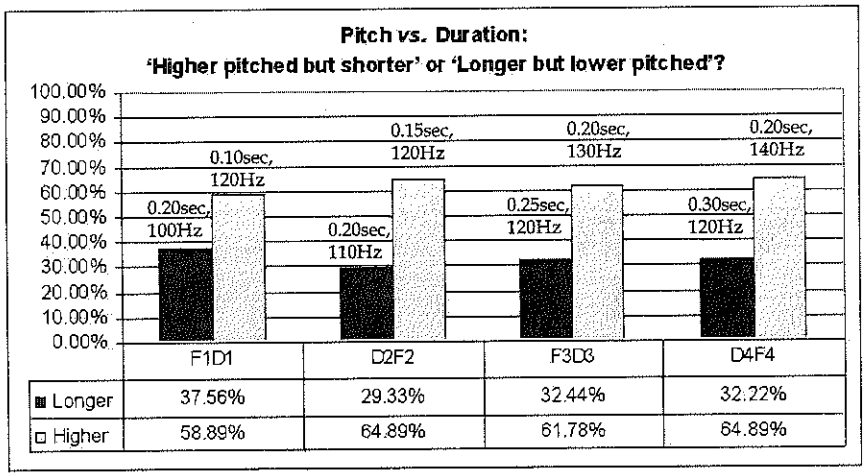


As can be seen from Figure 5, when the subjects are made to choose between a louder but lower pitched test word and a higher pitched but softer test word, the results are fairly consistent. Nearly 75% of the subjects choose a louder but lower pitched syllable over a softer but higher pitched syllable, in the cases *I1F1* and *F4I4*, when the difference in amplitude between the two vowels in the test words is 10dB. However, when the difference in amplitude is only 5dB, the percentage of the subjects making the same judgement drops below 65%, as in the case of *F2I2* and *I3F3*. This suggests that the subjects are sensitive to a larger difference in amplitude, and use that as a signal to prominence. On the other hand, when the difference in amplitude between two syllables is smaller, some subjects turn to using differences in F_0 as a signal to prominence – which explains a larger percentage of subjects picking the higher pitched syllable over the louder one as the more prominent item.

Pitch vs. Duration: 'Higher Pitched But Shorter' or 'Longer but Lower Pitched'?

In this section, comparison is made between a test word that is higher pitched and shorter, versus one that is lower pitched but longer. In the test word in which duration is manipulated, amplitude and fundamental frequency is kept constant at 70dB and 120Hz. In the test word in which the F_0 is being manipulated, amplitude and vowel length is kept constant at 70dB and 0.20sec. Figure 6 presents the results of the perception test for this set of manipulation.

Figure 6: Responses of subjects when choosing between higher pitched but shorter and longer but lower pitched syllables



From Figure 6, it can be observed that when the subjects are made to choose between F_0 and vowel length as perceptual cues, the results are very consistent. In every single step manipulation, they consistently choose the word containing the higher but shorter test word as the more prominent item, though the percentages, across all four manipulations, do not exceed 65%, suggesting that there is a rather large room for a different option.

Conclusion: A Hierarchy of Parameters

As can be seen, it is generally consistent that the SE subjects would use higher pitch, longer vowel length and increased loudness to determine prominence. When it comes to the relative strength of the perceptual cues for each group of subjects, a hierarchy of dominance of the parameters for prominence perception can be observed.

Intensity seems to be the dominant perceptual cue for prominence. The subjects would pick the test word containing the louder vowel as the more prominent item, even when the other item is higher pitched or longer. Intensity thus overrides pitch and duration as perceptual cues for prominence.

When the comparison is between a higher pitched and a longer test word, the choice is for the former. This group of subjects would pick the word containing the higher pitched vowel as the prominent item, as opposed to the longer one. Duration, is a perceptual cue, is the weakest, compared to intensity and pitch. Therefore, for the speakers of SE, in their perception of prominence, they would first use intensity, followed by pitch, and finally, duration.

The results have implications on previous research conducted on SE stress. As highlighted earlier, one of the assumptions researchers of stress in SE make is that the perceptual cues for prominence in SE is the same as that of BrE, namely, pitch. Researchers like Chua (1989) and Low (1998) readily assume that that a higher pitched syllable is a stressed syllable in their investigation of lexical stress placement patterns in SE, as how it is in BrE. Other researchers like Tongue (1974) and Platt and Weber (1980), in their experiments, use their own judgements, to determine the stressed syllables in their sample. It is to be noted of course that Tongue, Platt and Weber are not native speakers of SE. They use BrE perception to perceive prominence in SE, in which case, the higher pitched syllable is perceived to be the stressed syllable. Taking a higher pitched syllable as the stressed syllable, these researchers would have identified the 'incorrect' stressed syllables, and therefore would have come to a wrong conclusion about stress placement in SE. As the results from this paper show, the most dominant perceptual cue of prominence in SE is intensity, not pitch. The results presented in this paper have shown that these past research and their findings need to be re-examined, and that an analysis of stress without first determining the perceptual properties of prominence in could lead to wrong conclusions about stress placement in SE. It is hoped that future researchers no longer need to delve back into our colonial past in search of models, and that we could look forward to a more complete understanding of prominence perception in Singapore English, in its own right.

Appendix

Table 1: Values of the parameters manipulated for synthesis, manipulating 1 parameter at a time

Labels of the manip-ulations	V1 (the vowel in the 1st test word)			V2 (the vowel in the 2nd test word)		
F1	120 Hz	70 dB	0.20 sec	**100 Hz**	70 dB	0.20 sec
F2	120 Hz	70 dB	0.20 sec	**110 Hz**	70 dB	0.20 sec
F3	120 Hz	70 dB	0.20 sec	**130 Hz**	70 dB	0.20 sec
F4	120 Hz	70 dB	0.20 sec	**140 Hz**	70 dB	0.20 sec
I1	120 Hz	70 dB	0.20 sec	120 Hz	**60 dB**	0.20 sec
I2	120 Hz	70 dB	0.20 sec	120 Hz	**65 dB**	0.20 sec
I3	120 Hz	70 dB	0.20 sec	120 Hz	**75 dB**	0.20 sec
I4	120 Hz	70 dB	0.20 sec	120 Hz	**80 dB**	0.20 sec
D1	120 Hz	70 dB	0.20 sec	120 Hz	70 dB	**0.10 sec**
D2	120 Hz	70 dB	0.20 sec	120 Hz	70 dB	**0.15 sec**
D3	120 Hz	70 dB	0.20 sec	120 Hz	70 dB	**0.25 sec**
D4	120 Hz	70 dB	0.20 sec	120 Hz	70 dB	**0.30 sec**

Table 2: Values of the parameters manipulated for synthesis, manipulating 2 parameters at one time

Labels of the manip-ulations	V1			V2		
I1-F1	120 Hz	**60 dB**	0.20 sec	**100 Hz**	70 dB	0.20 sec
F2-I2	**110 Hz**	70 dB	0.20 sec	120 Hz	**65 dB**	0.20 sec
I3-F3	120 Hz	**75 dB**	0.20 sec	**130 Hz**	70 dB	0.20 sec
F4-I4	**140 Hz**	70 dB	0.20 sec	120 Hz	**80 dB**	0.20 sec
F1-D1	**100 Hz**	70 dB	0.20 sec	120 Hz	70 dB	**0.10 sec**
D2-F2	120 Hz	70 dB	**0.15 sec**	**110 Hz**	70 dB	0.20 sec
F3-D3	**130 Hz**	70 dB	0.20 sec	120 Hz	70 dB	**0.25 sec**
D4-F4	120 Hz	70 dB	**0.30 sec**	**140 Hz**	70 dB	0.20 sec
D1-I1	120 Hz	70 dB	**0.10 sec**	120 Hz	**60 dB**	0.20 sec
I2-D2	120 Hz	**65 dB**	0.20 sec	120 Hz	70 dB	**0.15 sec**
D3-I3	120 Hz	70 dB	**0.25 sec**	120 Hz	**75 dB**	0.20 sec
I4-D4	120 Hz	**80 dB**	0.20 sec	120 Hz	70 dB	**0.30 sec**

Note: The manipulation was kept as systematic as possible, with each step manipulation of each parameter paired with the same for another parameter, for example, *D1* (duration manipulated to 0.1 sec) in the first vowel and *F1* (F_0 manipulated to 100 Hz).

Notes

[1] Prominence is defined as the most perceptually salient syllable within a particular sentence/phrase compared to the other syllables within the same sentence or phrasal unit.

[2] In these studies, the stressed syllable is the one that receives the pitch obtrusion, following the British system of intonational analysis, identifying the head (the highest pitched syllable) as the stressed syllable within the intonational group.

[3] The Festival Speech Synthesis System: University of Edinburgh is a web-based synthesis tool that can be accessed online (www.cstr.ed.ac.uk/projects/festival/). The Festival Speech Synthesis System produces words based on the accents of several varieties of English in the United Kingdom as well as Spanish spoken in different regions.

[4] Praat is a speech-analysis software that can be accessed from http://www.praat.org.

[5] The average values for F0 in conversational speech in European languages are approximately 120 Hz for men (Fant, 1956). Normal conversation is conducted at about 70 dB (Moore, 1982:8).

[6] A test involving 6 listeners was run prior to the execution of this experiment to get at a set of values that, when two test items were compared, they were audibly distinguishable. Based on the responses of the 6 subjects, it was observed that a difference of 10Hz for F0, 5dB for amplitude and 0.05sec for vowel length were the smallest possible difference for the subjects to hear a distinction between two

[7] The just-noticeable difference for pitch perception is about 1 Hz, in the span of F0 from 80 to 160 Hz (Flanagan, 1957:534). The just-noticeable difference in intensity has a value of about 0.5 - 1 dB (Rodenburg, 1972) within the range of 20 dB to 100 dB (Miller, 1947). The just-noticable difference in duration between two sounds is about 10 - 40 msec (Lehiste, 1976:226), which is 0.01 - 0.04 sec. These values serve as an important starting point to which the variations in the parameters are decided and manipulated. However, as they are values for just-noticeable differences, the values chosen for the variations must invariably be larger than these differences, so as to ensure that the listeners would be able to clearly distinguish the differences between the sounds they hear.

[8] See Appendix Table 1 for the complete listing of the manipulations on F0, amplitude and length of the vowels.

[9] See Appendix Table 2 for the complete listing of the manipulations on F_0, amplitude and length of the vowels – when making comparison between the comparative strengths of the parameters.

[10] The 150 subjects consisted of 50 Chinese, 50 Malays and 50 Indians. A much detailed analysis on the ethnic differentiation in the perception of prominence can be found in Tan (2002).

[11] All 150 subjects who took the perception test are Singaporean, and never lived abroad for more than 5 years. They are all bilingual in both English and their respective Mother Tongues. For the Chinese subjects, besides Mandarin, some can also speak other Chinese languages like Teochew, Hokkien, and Cantonese. For the Indian subjects, only those who speak Tamil as their Mother Tongue were asked to participate in this experiment. All the Malay subjects have

Malay as their Mother Tongue. For all three groups of subjects, all of them use their respective Mother Tongue at least 50% of the time.

[12] The 'mother tongue' in the Singapore context, is not defined by Skutnabb-Kangas and Phillipson's criteria of origin, competence, function and identification. In Singapore, the Mother Tongue is the "superordinate language" (Gupta, 1998:117) of one's official ethnic group. The official languages of Mandarin, Malay and Tamil are assigned to the official ethnic groups correspondingly. Therefore, if one is ethnically classified as 'Chinese', then one's Mother Tongue is deemed to be Mandarin, that of a 'Malay', Malay and that of an Indian, 'Tamil' (1998:117).

References

Alsagoff, S. L. 1984. A Systematic Study of Word-accentuation Patterns in Singaporean English. Unpublished Academic Exercise. Department of English Language and Literature. National University of Singapore.

Awedyk, W. 1986. Perception of Stressed Syllables in Natural Stimuli: A Contrastive English-Polish Experimental Study. *Papers and Studies in Contrastive Linguistics* 21. pp. 41-46.

Bao, Z. M. 1998. The Sounds of Singapore English. In: J. A. Foley, *et al.* eds. *English in New Cultural Contexts: Reflections from Singapore.* Singapore: Oxford University Press. pp. 152-174.

Beckman, M. E. 1986. *Stress and Non-Stress Accent.* Dordrecht, The Netherlands: Foris Publications.

Beckman, M. E. and J. B. Pierrehumbert. 1986. Intonational Structure in English and Japanese. *Phonology Yearbook 3.* pp. 255-310.

Bolinger, D. L. 1958. A Theory of Pitch Accent in English. *Word* 14. pp.109-149.

Chua, Y. L. 1989. Some Factors Determining Word Stress in Educated Singapore English. Unpublished Academic Exercise. Department of English Language and Literature. National University of Singapore.

Deterding, D. H. 1994a. The Characteristics of Singapore English Pronunciation. *Review of Educational Research and Advances for Classroom Teachers.* 1.

Deterding, D. H. 1994b. The Intonation of Singapore English. *Journal of the International Phonetic Association* 24. 2:61-72.

Deterding, D. H. and R. Hvitfeldt. 1994. The features of Singapore English Pronunciation: Implications for Teachers. *Teaching and Learning.* National Institute of Education, Singapore 15. 1:98-107.

Eek, A. 1987. The Perception of Word Stress: A Comparison of Estonian and Russian. In: R. Channon, and L. Shockey. eds. *In Honor of Ilse Lehiste.* Dordrecht, The Netherlands: Foris Publications. pp. 19-32.

Fant, G. 1956. On the Predictability of Formant Levels and Spectrum Envelopes from Formant Frequencies. In: M. Halle, H. Lunt and H. MacLean. eds. *For Roman Jakobson.* The Hague: Mouton. pp. 109-120.

Flanagan, J. L. 1957. Estimates of the Maximum Precision Necessary in Quantizing Certain 'Dimensions' of Vowel Sounds. *Journal of the Acoustical Society of America* 24. 533-534.

Fry, D. B. 1955. Duration and Intensity as Physical Correlates of Linguistic Stress. *Journal of the Acoustical Society of America* 27. 4:765-768.

Fry, D. B. 1958. Experiments in the Perception of Stress. *Language and Speech* 1. 126-152.

Fry, D. B. 1965. The Dependence of Stress Judgements on Vowel Formant Structure. In: E. Zwirner and W. Bethge. eds. *Proceedings of the Sixth International Congress of Phonetic Sciences, Münster, 1964.* Basel: Karger. pp. 306-311.

Hasegawa, Y. and K. Hata. 1992. Fundamental Frequency as an Acoustic Cue to Accent Perception. *Language and Speech* 35. pp. 87-98.

Janota, P. 1979. Some Observation on the Perception of Stress in Czech. *Proceedings of the Ninth International Congress of Phonetic Sciences.* Copenhagen: B. Stougaard Jensen. pp. 403-409.

Jassem, W. 1959. The Phonology of Polish Stress. *Word* 15. pp. 252-269.

Jassem, W., J. Morton and M. Steffen-Batóg. 1968. The Perception of Stress in Synthetic Speech-like Stimuli by Polish Listeners. *Speech Analysis and Synthesis* 1. pp. 289-308.

Kachru, B. 1979. Models of English for the Third World: White Man's Linguistic Burden or Language Pragmatics? In: J. Richards. ed. *New Varieties of English: Issues and Approaches.* Occasional Papers No. 8. Singapore: SEAMEO Regional Language Centre.

Ladefoged, P. 1967. *Three Areas of Experimental Phonetics.* London: Oxford University Press.

Lehiste, I. 1976. Suprasegmental Features of Speech. In: Norman J. Lass, ed. *Contemporary Issues in Experimental Phonetics.* New York: Academic Press. pp. 225-239.

Low, E. L. 1994. Intonation Patterns in Singapore English. Unpublished M.Phil. Dissertation. Department of Linguistics, University of Cambridge.

Low, E. L. 1998. Prosodic Prominence in Singapore English. Unpublished P.h.D. Dissertation. University of Cambridge.

Low, E. L. and E. Grabe. 1999. A Contrastive Study of Prosody and Lexical Stress Placement in Singapore English and British English. *Language and Speech* 42. 1:39-56.

Miller, G.A. 1947. Sensitivity to Changes in the Intensity of White Noise and Its Relation to Masking and Loudness. *Journal of the Acoustical Society of America* 19. 609-619.

Mitsuya, F. and M. Sugito. 1978. A Study of the Accentual Effect on Segmental and Moraic Duration in Japanese. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics* 12. 97-112.

Mohanan, K. P. 1992. Describing the Phonology of Non-native Varieties of a Language. *World Englishes* 11. pp. 111-128.

Moore, B. C. J. 1982. *An Introduction to the Psychology of Hearing.* 2nd edition. Orlando FL: Academic Press.

Morton, J. and W. Jassem. 1965. Acoustic Correlates of Stress. *Language and Speech* 8. pp. 159-181.

Ng, N. Y. 1985. Word Stress in Educated Singaporean English. Unpublished Academic Exercise. Department of English Language and Literature. National University of Singapore.

Platt, J. and H. Weber. 1980. *English in Singapore and Malaysia.* Kuala Lumpur, New York: Oxford University Press.

Rodenburg, M. 1972. Sensitivity of the Auditory System to Differences in Intensity. Unpublished P.h.D. Dissertation. Medical Faculty. Rotterdam.

Sng, S. Y. B. 1991. Word Accent in Singapore English. Unpublished Academic Exercise. Department of English Language and Literature. National University of Singapore.

Studdert-Kennedy, M. 1976. Speech Perception. In: N. J. Lass, ed. *Contemporary Issues in Experimental Phonetics:* New York: Academic Press. pp. 243-293.

Tay, M. W. J. 1982. The Phonology of Educated Singapore English. *English World Wide* 3. 2:135-145.

Tongue, R. K. 1979. *The English of Singapore and Malaysia.* 2nd edition. Singapore: Eastern Universities Press.

Weitzman, R. 1969. Japanese Accent: An Analysis Based on Acoustic-phonetic Data. Unpublished P.h.D. Dissertation. University of Southern California.

Westin, K., R. G. Buddenhagen and D. H. Obrecht. 1966. An Experimental Analysis of the Relative Importance of Pitch, Quantity, and Intensity as Cues to Phonemic Distinctions in Southern Swedish. *Language and Speech* 9. pp.114-126.

10

ADJACENCY PAIRS: QUESTIONS AND ANSWERS IN INTERVIEWS

Kulwindr Kaur

According to Richards, Platt and Platt (1992:7) an **adjacency pair** is "a sequence of two related utterances by two different speakers. The second utterance is always in response to the first". They suggest that adjacency pairs are part of the structure of conversation and are studied in conversational analysis. An adjacency pair is sequenced into two parts, commonly known as the *first and second pair parts* and is regarded by conversation analysts as the basic or fundamental unit in conversation (Coulthard, 1977). Examples of adjacency pairs include complaint-denial, question-answer, greeting-greeting and invitation-acceptance/rejection, offer-acceptance/non-acceptance, complaint-apology sequences (Richards, Platt and Platt, 1992:7).

For this study a question will be defined as an illocutionary act (Searle, 1969:66 cited in Pillai, 1996:13) which seeks "to elicit verbal information" from the other party. On the other hand, an answer is defined as a verbal response to the preceding question. A question and answer sequence fulfills the conditions of adjacency pair in that it consists of two utterances by two different speakers. A question utterance constitutes the first pair part while an answer is the second pair part. As Sacks (1969, quoted in Coulthard and Brazil, 1992:52, also cited in Pillai, 1996:40) observes, "given a question, regularly enough an answer will follow".

For this study, an utterance is defined as "a sentence, or sometimes strings of sentences paired with a context" (Levinson, 1983:19 cited in Pillai, 1996:16). In relation to the two sentences in a question-answer adjacency pair, this means that in discourse what is said by any one person before or