# TOWARDS A PHONOLOGY OF SINGAPORE ENGLISH

**Tony T.N. Hung**
National University of Singapore

## 1. INTRODUCTION

At the present state of research, almost any paper on the phonology of Singapore English has to begin with a number of caveats, and this is no exception.

Firstly, it has to be pointed out that I am not yet in a position to make any statistically valid claims about educated Singaporean English speakers in general. I have been investigating the speech of an individual, an undergraduate at the National University of Singapore, whom I judged to be representative of the younger generation of educated Singaporean speakers who use English as their primary language -- which is the type of speakers I am most interested in. The investigation of the linguistic systems of individuals is both a valid enterprise in its own right and a necessary first step towards a general description of a language or variety of language.

Secondly, a crucial distinction needs to be made between what speakers do and what they know. For example, the majority of educated Singaporean speakers can hear the difference between RP *sit* [sɪt] and *seat* [siːt], and may even be able to mimic these pronunciations, but in their own spontaneous, natural speech, no distinction is normally made, both words being pronounced [sɪt]. (Incidentally, a parallel situation exists in Singapore Mandarin. Most educated Singaporean speakers of Mandarin know that there is a distinction in Beijing Mandarin between dental, palatal and retroflex obstruents, as illustrated by the initial consonants in *si* ('four'), *xi* ('west'), and *shi* ('matter') respectively. But in their own speech, there is no such distinction, all three being pronounced with a dental or alveolar [s].

What I am investigating in this paper is what speakers actually do, not what they know. It is quite possible for some Singaporeans to think that their pronunciation is different from what I am describing, and that they make certain distinctions, when in actual fact they do not in their natural -- as opposed to artificial and self-conscious -- speech.

Thirdly, the phonological system of a speaker is a complex and multi-faceted thing, and is far more than a mere inventory of sounds. In the context of this paper,

I am naturally not in a position to describe the phonological system of even an individual speaker in any comprehensive manner, but only certain salient aspects of it. Though, ultimately, a linguistic system ought to be analysed on its own terms, in the present paper I shall be making references to other well-known varieties of English such as British RP or General American (GA) as a time-saving expedient by focusing on their differences from Singapore English (SE), without any implication that one is 'correct' and the other 'deviant'.

Indebted as I certainly am to previous work in the field (some of which are listed in the Bibliography), I shall not have time to support or disagree with their various analyses, and I shall proceed with as few prior assumptions as possible. As far as possible too, I have resorted to instrumental data to provide objective support for my own analysis, something which few previous investigators have done.

Describing the phonology or sound system of a language (or, more precisely, dialect of a language) involves a number of components, including:

(1) An inventory of PHONEMES, or sound segments which contrast with each other;

(2) Conditions governing how segments combine or co-occur with other segments, i.e. the DISTRIBUTION of segments;

(3) Regular and predictable variations that are conditioned by the environment, i.e. ALTERNATION.

In addition, there are higher-level phenomena such as stress, rhythm and intonation, but these are beyond the scope of the present paper.

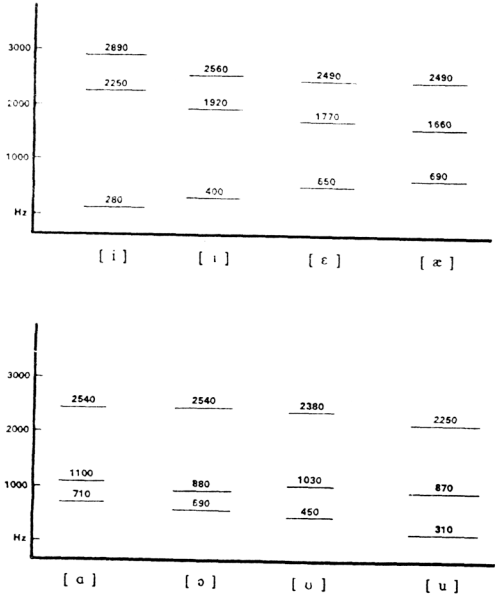
## 2. THE VOWEL SYSTEM

One of the major differences between Singapore English and other Englishes lies in their vowel systems. This is not surprising, considering that (unlike consonants) vowels do not involve any contact between two articulators, but are differentiated by various degrees of tongue height and frontness, and by lip-rounding. The tongue may be advanced and raised, as in [i], or lowered, as in [æ], or retracted and raised, as in [u], or lowered, as in [ɑ], but there is an almost infinite number of intermediate possibilities. This is why, more than anything else, vowels set one accent apart from another, and even within the same speech community, vowel qualities fluctuate to some extent from individual to individual.

Though we cannot easily see the position of the tongue in the mouth, we can rely on instrumental measurements, in particular spectrograms, to capture the qualitative differences between vowels. For the sake of those who may not be familiar with its technicalities, let me provide a simple explanation.

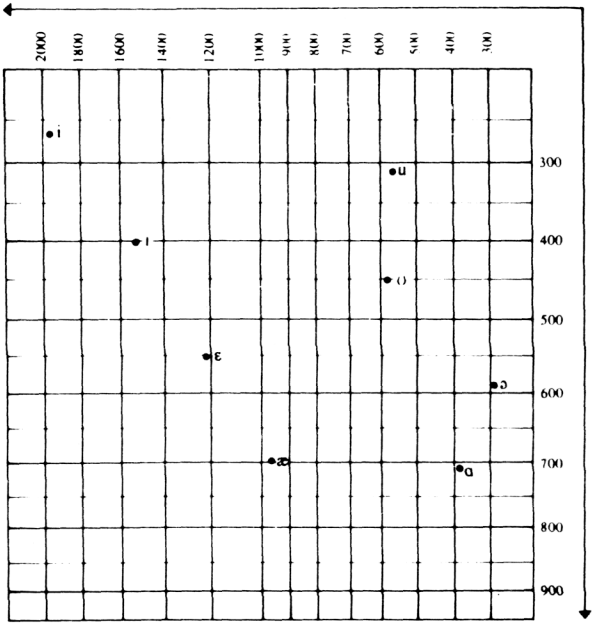
Because of the complex shape of the oral cavity in the production of each vowel, caused by various tongue and lip positions, the air resonates at different frequencies all at once, which show up as dark bands of concentrated energy at various frequencies on a spectrogram. The lowest of these is called the First Formant or F1, the Second F2, etc. What is relevant for our purposes is that there is an inverse correlation between F1 and the height of the tongue, namely: the higher the F1, the lower the tongue position, and vice versa. There is also a correlation between the distance between F1 and F2 (call it F2') and the backness of the tongue, namely: the smaller the distance, the more back the tongue is, and vice versa. As an illustration, the frequencies of the first three formants of vowels in General American are given in Figure 1 below (notice that the differences are of the order of over 100 Hz).

FIGURE 1: FORMANT FREQUENCIES OF AMERICAN VOWELS (from Ladefoged 1993)



When F1 is plotted on the vertical axis of a graph, against F2' on the horizontal axis, with zero at the top right hand corner, the configuration of the vowels resembles what is found in traditional vowel charts, which are arranged according to vowel height and backness (see Figure 2).

FIGURE 2: FORMANT CHART OF AMERICAN VOWELS




A spectrographic analysis of the vowels of my Singaporean subject yields the picture given in Figure 3. I have grouped the data on the basis of how their F1 and F2 measurements cluster together, without any preconceptions about possible phonemic contrasts.

FIGURE 3:
FORMANT FREQUENCIES OF VOWELS OF A SINGAPORE ENGLISH SPEAKER

| WORD | VOWEL F1 Hz | F2 Hz | DUR Sec | WORD | VOWEL F1 Hz | F2 Hz | DUR Sec |
|---|---|---|---|---|---|---|---|
| Set 1 | | | | | | | |
| 1. this | 260 | 2120 | .104 | 2. these | 260 | 2160 | .131 |
| 3. bit | 260 | 2240 | .106 | 4. beat | 260 | 2220 | .101 |
| 5. dim | 260 | 2240 | .150 | 6. deem | 280 | 2260 | .201 |
| MEAN: | 263 | 2207 | .132 | | | | |
| Set 2 | | | | | | | |
| 7. bend | 520 | 1780 | .179 | 8. band | 540 | 1880 | .209 |
| 9. bet | 540 | 1860 | .131 | 10. bat | 520 | 1800 | .106 |
| 11. bared | 520 | 1800 | .180 | 12. bad | 520 | 1820 | .195 |
| MEAN: | 526 | 1823 | .166 | | | | |
| Set 3 | | | | | | | |
| 13. hairs | 700 | 2120 | .318 | 14. has | 720 | 2060 | .192 |
| 15. at | 680 | 2060 | .126 | 16. act | 720 | 1900 | .106 |
| MEAN: | 705 | 2035 | .185 | | | | |
| Set 4 | | | | | | | |
| 17. full | 240 | 880 | .284 | 18. fool | 240 | 840 | .225 |
| 19. foot | 300 | 1000 | .101 | 20. food | 300 | 940 | .095 |
| MEAN: | 256 | 915 | .176 | | | | |
| Set 5 | | | | | | | |
| 21. buck | 660 | 1180 | .110 | 22. bark | 720 | 1220 | .100 |
| 23. luck | 680 | 1200 | .087 | 24. lark | 700 | 1220 | .093 |
| 25. lust | 680 | 1340 | .128 | 26. last | 780 | 1420 | .151 |
| MEAN: | 673 | 1240 | .108 | | 733 | 1286 | .114 |
| Set 6 | | | | | | | |
| 27. lord | 520 | 1040 | .164 | 28. lot | 620 | 1120 | .112 |
| 29. jaws | 520 | 1020 | .243 | 30. joss | 520 | 1160 | .218 |
| 31. caught | 560 | 940 | .128 | 32. cot | 580 | 960 | .103 |
| 33. door | 520 | 1140 | .209 | | | | |
| MEAN: | 548 | 1054 | .168 | | | | |

Set 7

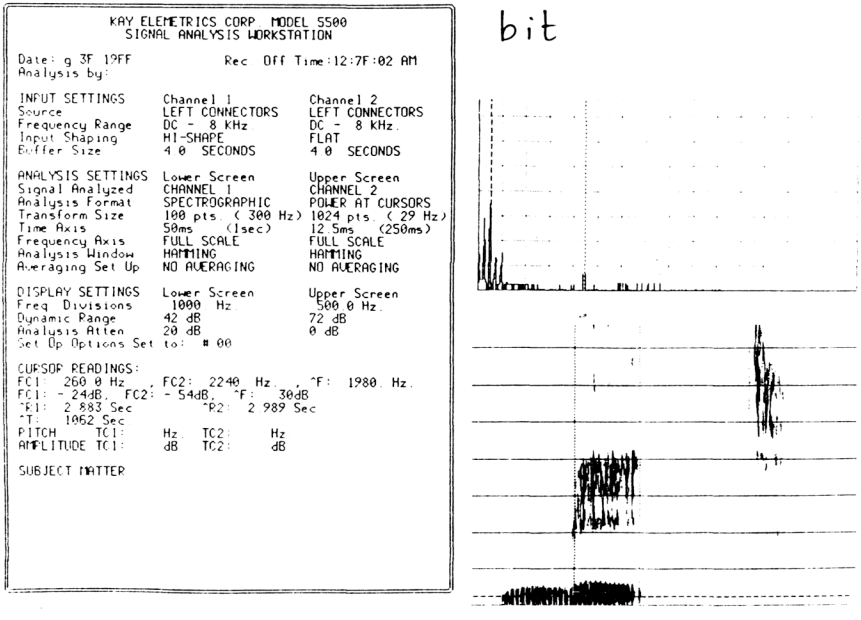
| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 34. day | 400 | 2200 | .268 | 35. date | 400 | 2140 | .135 |
| 36. pay | 440 | 2380 | .271 | 37. take | 460 | 2140 | .103 |
| MEAN: | 425 | 2215 | .194 | | | | |

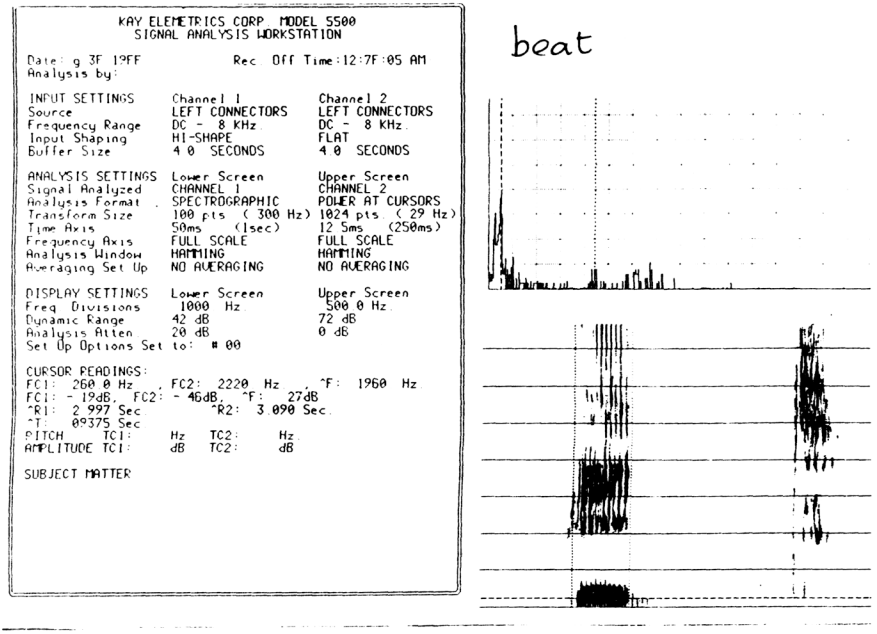
Set 8

| | | | |
|---|---|---|---|
| 38. doe | 500 | 1000 | .206 |
| 39. toe | 400 | 900 | .228 |
| 40. coat | 420 | 960 | .106 |
| MEAN: | 440 | 953 | .180 |

Certain tentative observations and conclusions can be drawn from the data in Figure 3. Firstly, the number of vowel contrasts in my subject's English is smaller than in RP or General American. In particular, the vowels in pairs of words like *bit beat*, *bet bat*, *foot food* and *cot caught* are not distinctive, either in their formant frequencies or their duration. For illustration, a spectrographic printout for *bit beat* as spoken by my informant is given in Figure 4.

FIGURE 4: SPECTROGRAMS FOR *BIT, BEAT*:



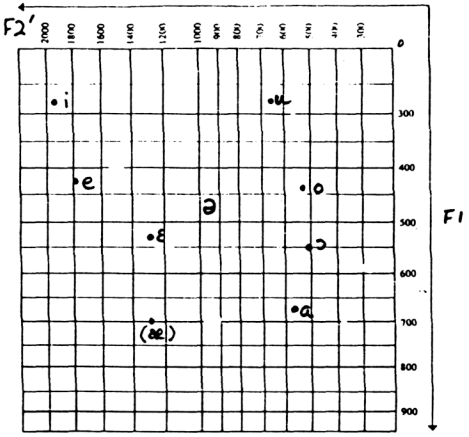



Sets 2 and 3 of the data in Figure 3 pose an interesting puzzle. The acoustic properties of the vowels of Set 3 are sufficiently different from Set 2 to merit their being treated as potential members of a different phoneme, unless they are predictable variants of Set 2, which remains to be seen. It is also not easy to decide whether Set 5 contains 1 or 2 vowel phonemes, /a/ and /ʌ/.

It can be noticed also that the vowels in words like *day, doe, care* are monophthongs rather than diphthongs. Their spectrograms show a (more or less) flat F1 and F2 instead of a diverging or converging one. There are 5 true diphthongs in SE, in words like *tie, toy, cow, beer, poor* respectively.

Tentatively then, we may postulate the vowel chart given in Figure 5 for my Singaporean subject.

FIGURE 5: VOWEL CHART FOR SINGAPOREAN SPEAKER




The total inventory of vowel phonemes for this speaker is summarised below:

---

VOWEL PHONEMES

| VOWEL | EXAMPLE |
|---|---|
| i | bit, beat |
| e | day, late |
| ɛ | bet, bat, |
| ?(æ | hair, act) |
| u | fool, full |
| o | toe, load |
| ɔ | cot, caught |
| a | luck, lark |
| ə | fur, bird |
| ai | tie, high |
| ɔi | toy, boil |
| *au* | *cow, town* |
| iə | beer, here |
| uə | poor, sure |

---

## 3. CONSONANTS

By and large, SE has the same inventory of consonantal phonemes as RP and GA. In SE (as in several other varieties of English), the initial consonants in words like *thin* and *this* are realised as the plosives [t] and [d] respectively, an impression supported by corresponding blank spaces on a spectrogram (evidence for plosives, in which the airstream is completely blocked). In isolation, pairs of words like *thin tin* or *though doe* are virtually indistinguishable, and subjects asked to identify the words have shown about 50% chance of success.

Does this mean that what are represented distinctively as /θ/ and /ð/ in RP and GA are actually the same phoneme as /t/ and /d/ respectively in SE? Matters are not so simple here. There is evidence that, even in SE, /θ/ and /ð/ (however one may choose to represent them) do not belong to the same phoneme as /t/ and /d/ respectively, as there are environments in which they are realised differently, e.g. *health* [hɛlf] vs. *healthy* [hɛlti], as opposed to *guilt* [gilt] vs. *guilty* [gilti]. It is therefore quite possible that there is a separate phoneme in SE (represented in other accents as /θ/) which is distinct from /t/ and /f/, and which is phonetically realised as [t] in the onset and [f] in the coda of a syllable. Obviously, further data and analysis are required before any such conclusion can be drawn.

## 4. DISTRIBUTION

The distribution of sound segments in a language can best be stated in terms of the syllable. Every language has its own constraints on what types of segments, and how many of them, can occur in which parts of the syllable: at the beginning (onset), center (nucleus) or ending (coda).

The syllable in SE allows a maximum of 3 consonants in the onset, as in *strike* and *spring*, subject to the same phonotactic constraints as in RP or GA. In the coda, however, the possibilities are more limited. Whereas RP and GA allow up to 4 consonants, as in [tɛksts] (*texts*) and [glimpst] (*glimpsed*), for most SE speakers the upper limit seems to be either 2 or 3, such that the above words are pronounced as [tɛks] and [glims]/[glimst] respectively.

Apparently as a consequence of these syllable-structure constraints, final consonant clusters are regularly simplified in SE, by the deletion of some of the word-final consonants. Certain types of consonants, e.g. plosives, are more liable to be

deleted, and certain environments, i.e. when preceded by a homorganic consonant, are more liable to trigger deletion. Some of the most typical patterns of deletion are as follows:

(i) Voiceless plosives are deleted in word-final clusters if preceded by another voiceless consonant, e.g. [as] *ask*, [fis] *fist*, [sɔf] *soft*, [luk] *looked*, [las] *last*, etc. On the other hand, they are not normally deleted after voiced segments, e.g. [ant] *aunt*, [dont] *don't*, [bilt] *built*, etc.

(ii) Voiced plosives are deleted if preceded by a homorganic consonant in word-final clusters, e.g. [frɛns] *friends*, [ɛn] *end*, [graun] *ground*.

(iii) /l/ is often deleted if (a) preceded by a [-front] vowel in a word-final cluster, e.g. [trɛvə] *travel*, [o] *old*, [wɔ] *wall*, as opposed to [ail] *I'll*, [wil] *will*, [wɛl] *well*, etc., where /l/ is preceded by a [+front] vowel; (b) adjacent to a nasal consonant, e.g. [oni] *only*, [ɛm] *elm*.

There are certain combinations of consonants in the coda which are possible in RP and GA but not possible with most SE speakers, particularly [lm] and [sp]. When these sequences occur in a word, e.g. in *film* and *wasp*, a repair strategy comes into play to switch the two sounds, in one case to [flim] and in the other [waps].

Consonants rarely or never occur in the nucleus of a syllable in SE, in contrast with syllabic consonants in RP and GA in such words as [gɑdn] *garden*, [lɪtl] *little* and [prɪzn] *prison*; for these words, [gadən], [litəl] and [prizən] are the typical SE pronunciations.

## 5. ALTERNATION

In most languages, it is a common phenomenon for sounds to undergo modification or alternation in different environments. Some of these are phonetically motivated, e.g. the lip-rounding of consonants preceding rounded vowels ([t] and [k] in *too* and *cool*), but others are not so motivated, and are specific to a particular language or dialect. In SE, a number of these phonological alternations can be noted, of which I can mention only a few examples here:

(i) Devoicing: There is a tendency to completely devoice underlyingly voiced obstruents in word-final position, e.g. [bilif] *believe*, [dʒʌtʃ] *judge*, [rip] *rib*. In other varieties of English, word-final obstruents are also partially devoiced, but not as

completely as in SE, where the /dʒ/ in *judge* is as voiceless as the /tʃ/ in *batch*. The opposite tendency, to voice a voiceless consonant in intervocalic position, e.g. [diz is] (*this is*) and [prɛʒə] (*pressure*), seems to be somewhat sporadic, and limited to alveolar fricatives.

(ii) Word-final plosives are usually unreleased in SE.

(iii) Glottal stop insertion: This is a particularly variable phenomenon (in this as in other dialects of English). In SE, [ʔ] is inserted mainly in two environments, namely (a) before a word-initial vowel (especially when the preceding word ends in a vowel), and (b) before a word-final stop (which may subsequently get deleted), as in [meʔk] *make*, [laiʔ] *like*.

Some of the phonological processes which are common in RP and GA are not commonly found in SE, e.g.:

(i) Place assimilation: e.g., in words like [inkəm] *income* and [gudbai] *goodbye*, the |n| and |d| are seldom velarised or labialised (to |ŋ| and |b|) respectively in SE;

(ii) There is in SE little if any variation in vowel length which is due to the phonetic environment. E.g. vowels are not appreciably lengthened before voiced segments. There is, however, a tendency in SE to lengthen vowels in word-final and phrase-final syllables (e.g. [hɛpi.] *happy* vs. [hɛpinə.s] *happiness*).

(iii) Vowel reduction: Reduction of unstressed vowels to |ə|, especially in positions preceding a stressed syllable, is not common in SE. E.g., the first syllables in *familiar, conclusion, official* etc. retain their full values instead of being reduced to |ə|

## REFERENCES

Brown, A. 1988. Vowel differences between Received Pronunciation and the English of Malaysia and Singapore. In J. Foley (ed.), New Englishes: The Case of Singapore. Singapore University Press.

Ladefoged, P. 1993. A Course in Phonetics, 3rd Ed. Harcourt.

Lim, H. 1992. A Spectrographic Analysis of Vowels in Singapore English. Academic Exercise, Dept. of English Language and Literature, National University of Singapore.

Platt, J. & Weber, H. 1980. English in Singapore and Malaysia. Oxford University Press.

Tay, M., & Gupta, A.F. 1981. Towards a description of Standard Singapore English. In R.B. Noss (ed.), Varieties of English in Southeast Asia. RELC.

Tay, M. 1982. Phonology of educated Singapore English. In English World-Wide, 3:2:135-145.

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

Wells, J.C. 1982. Accents of English (3 volumes). Cambridge University Press.