Malaysian English: an instrumental analysis of vowel contrasts

STEFANIE PILLAI*, ZURAIDAH MOHD. DON**, GERALD KNOWLES***, and JENNIFER TANG****

ABSTRACT: This paper makes an instrumental analysis of English vowel monophthongs produced by 47 female Malaysian speakers. The focus is on the distribution of Malaysian English vowels in the vowel space, and the extent to which there is phonetic contrast between traditionally paired vowels. The results indicate that, like neighbouring varieties of English, Malaysian English vowels occupy a smaller vowel space than those of British English. The lack of contrast in vowel quality between vowel pairs was more apparent for /i/-/I/, /e/-/æ/ and /æ/-/ə/. However, there was a considerable difference in duration in the vowel pairs except for /o/-/ɔ/ which contrasted in terms of vowel quality.

INTRODUCTION

The kind of English used in Malaysia has undergone substantial change from the time of its introduction in the 18th century. English in Malaysia has many varieties, ranging from the more colloquial to the standard form of English, spoken in many different local accents, and used in a variety of social and professional contexts. Although the term Malaysian English (MalE) is sometimes taken to refer to the colloquial variety rather derogatively called “Manglish”, it is used here as an umbrella term to include all varieties of English used in Malaysia (see Gaudart 1997; Morais 2001). Malaysian English is indeed a complex entity used by Malaysians of different ethnic, geographical and educational backgrounds, with different levels of proficiency. A small minority of Malaysians uses English as a first language (Crystal 1997); while for others it may range from being a dominant home language to not being used at all at home. The extent to which English is used in everyday communication also depends on factors like geographical location and profession. All these variables have resulted in the various sub-varieties of Malaysian English being coloured by different accents, vocabulary and morphological, and syntactic variation (e.g. Baskaran 2004; 2005).

Descriptions of Malaysian English, however, have tended to simplify the distinction among the various sub-varieties, into either two or three categories (e.g. Platt and Weber 1980; Baskaran 1994). In general, there is posited to be an acrolectal variety which does not differ much in morpho-syntactic variation from Standard English, and a more colloquial variety of Malaysian English with considerably more phonological, lexical and
morpho-syntactic variation. However, as expected, such a simplistic division ignores the array of linguistic differences arising from the socio-cultural factors mentioned above. Perhaps because of the association of MalE with the colloquial variety and also because this sub-variety of MalE is more likely to exhibit more obvious ‘differences’, current research on MalE tends to focus on it. In terms of pronunciation, the standard variety of MalE is said to be similar to received pronunciation (RP), in which respect it resembles other acrolects and is atypical of MalE; what is referred to as typical MalE pronunciation is generally based on an impressionistic analysis of colloquial MalE, which tends to be used with more ethnically “marked” (Gill 2002: 64) accents (Baskaran 1994; Rajadurai 2006). There is to date a dearth of research based on the instrumental analysis of sounds, and this is the gap that the study reported in this paper seeks to fill.

BACKGROUND

Descriptions of MalE vowels generally refer to a lack of distinction between the traditionally paired vowels, and a tendency to produce monophthongs in place of RP diphthongs (Wong 1984; Nair-Venugopal 2003; Rajadurai 2004; Baskaran 2005). For instance, Platt and Weber (1980: 172) describe the diphthong in so as similar to the RP vowel in hawk and caught but shorter. Similarly the /es/ diphthong is said to be produced similar to the RP [ɛː] or [ɛ] (Platt and Weber 1980: 172–3). The lack of distinction between paired vowels can be attributed to the tendency to shorten long vowels (Platt and Weber 1980; Zuraidah 1997; Baskaran 2005). Word pairs such as cart and cut, cot and caught are consequently likely to be homophones in MalE. The status of such words as homophones can also be due to vowel quality, since vowel quality in MalE is different from RP. Baskaran (2005: 28), for example, describes /ɔ/ and /ɒ/ as being produced as a “half-open and a more central vowel” in MalE. The former is similar to the findings reported in an instrumental study of vowels by Pillai, Knowles and Zuraidah (forthcoming) in which the vowel in the word north was produced more open and front in MalE than by British respondents (Deterding 1997). The other two vowels reported in this study, namely the vowels of wind and sun, were also produced differently than in British English, both being produced slightly more open and front than British English vowels. The findings were based on recordings of 15 fluent speakers who were recorded reading the North Wind and the Sun text. Differences in vowel quality and the lack of a length distinction result in a smaller vowel system than RP (Zuraidah 1997; Pillai, Knowles and Zuraidah 2006; Pillai et al. forthcoming).

Similar findings have been reported for Singapore and Brunei English, which is not surprising given the geographical proximity of these two countries to Malaysia. These are also appropriate comparator varieties in view of the fact that they share with Malaysian English, a common historical origin which links them all to modern RPs. For instance, Deterding (2003: 6–7) noted that the vowel space for the pairs /ɪ/ and /I/, /e/ and /æ/, /ɔ/ and /ɒ/, were closer together in Singapore English than in British English. Similar findings were reported for Hong Kong English (Deterding, Wong and Kirkpatrick 2008: 162). On the other hand, Deterding (2003: 8) reported that /u/ and /ʊ/ showed a greater contrast in Singapore English, with /u/ being made further back than in British English. The respondents in both studies were undergraduates of Chinese ethnicity who were considered by the researchers to be fluent speakers of English.

In a similar study of Brunei English, Salbrina (2006), also found a lack of distinction in vowel duration which resulted in the conflation of long and short vowels. Her subjects
consisted of 10 female Malay speakers of Brunei English, who were recorded reading the *North Wind and the Sun* passage. As in the case of Singapore English, Salbrina (2006: 254), found that the vowel pairs /i:/ and /i/, /ɔ:/ and /o/, and /e/ and /æ/ were not distinguished. She also reported the same findings for /u:/ and /o/, which like Hong Kong English (Deterding et al. 2008: 163) were more frontal than in Singapore English. In the case of the diphthongs, Salbrina (2006: 263) found evidence of monophthongization as in the case of the /eɪ/ diphthong, which was produced as a long /eː/.

The reason for the different realizations of English vowels is often attributed to the influence of speakers’ first language (L1). For example, Malay has a smaller vowel inventory and no vowel length distinction (Soo 1990; Zuraidah 1997; Nair-Venugopal 2000), and consequently the vowels of Malay speakers “tend to have equal duration, and RP vowels which do not exist in Standard Malay tend to be pronounced with Standard Malay vowel quality” (Zuraidah 1997: 38). She found that the speakers in her study, which consisted of 12 native speakers of Malay aged between 20 and 23, seemed to produce a different vowel because of differences in vowel quality and quantity between Malay and English. The following pairs of vowels were realized as single vowels (Zuraidah 1997: 38–40)

- [i] and [ɪ] become [i]
- [u] and [ʊ] become [u]
- [ɛ] and [æ] become [ɛ]
- [o] and [ɔ] become [o]
- [ʌ] and [æ] become [a]
- [ɔ] and [ɔ] become [ɔ]

However, another study of Malay speakers of English found that the paired vowels were indeed contrasted, albeit not as sharply as, for example, in British English (Subramaniam 2008). The different results in the two studies could be due to differences in the subjects’ level of English proficiency, and different methods of data collection and analysis. Whereas one was based on auditory impressions of the audio recordings (word and utterance level) of 12 Malay undergraduates (Zuraidah 1997), the other was based on an instrumental analysis of recordings of informal interviews with five Malays considered to be fluent speakers of English (Subramaniam 2008).

We have to treat with some caution claims concerning the extent to which the L1 of Malaysian speakers influences their English pronunciation, since we are dealing with a heterogeneous group with a wide range of L1s, and ethnic, geographical, educational and socio-economic backgrounds. Valid generalizations about vowels produced by Malaysians cannot be made using only Malay as a reference point, particularly when referring to Malaysian English in general (for example, see Baskaran 2005). Further, it has been shown that Malaysians practise accent-switching, and can move from a more to a less ethnically marked accent depending on the speaking context (see Gill 2002; Rajadurai 2004; Pillai and Fauziah 2006). Some speakers, for example, tend to use less marked features when they switch to the acrolectal variety of MalE, and so it is not easy to perceive their ethnicity by just listening to them (Pillai et al. forthcoming).

**THE PRESENT STUDY**

As previously mentioned, most current descriptions of MalE pronunciation tend to be based on auditory impressions, and there is therefore a need to conduct a more systematic
analysis. Without such an analysis, there is a danger that the description given of MalE pronunciation may not be accurate, and could amount to a stereotypical description, which will not contribute to the identification of local or regional norms. With this in mind, the Corpus of Malaysian English (COMEL) project was initiated to build a database of spoken MalE to be analysed at various levels. This paper reports the findings of a study undertaken to analyse monophthongs in MalE, and sets out to answer the following research questions in relation to the monophthongs of MalE:

1. What is the distribution of vowels in MalE in the vowel space as determined by the frequencies of the first and second formants of the vowels?
2. To what extent is there a contrast in MalE between the vowels traditionally paired in terms of vowel quality and duration?

METHODS

In order to address the two research questions, instances of 11 English monophthongs produced by 47 female speakers of MalE were analysed instrumentally.

Data

The data were taken from recordings made by 47 female Malaysian undergraduates from different ethnic backgrounds: 31 Chinese, 10 Indians, 5 Malays and 1 Eurasian. Twelve of the students cited English as their first language (L1): 9 Chinese, 2 Indians and 1 Eurasian. Among the students of Chinese ethnicity, 12 reported Mandarin as their L1, while the others spoke other Chinese varieties: Cantonese (5), Hokkien (4) and Hakka (1). Most of the students of Indian origin had Tamil as their L1 (7), while only one student had Malayalam as her L1. All the Malay respondents had Malay as their L1.

Since the study sought to determine the pronunciation of vowels of fluent speakers of English, the students selected were all English language majors who could be assumed to be proficient in English. Only recordings from female students were selected for analysis in order to keep the gender variable consistent, and this was done because there were too few recordings made by male students to make a valid comparison. Also in view of pitch differences in male and female voices, it was considered better to avoid analysing male and female recordings together.

The target vowels were embedded in a CVC context (where C = a stop consonant) and placed in a carrier sentence: Please say CVC again. Respondents were presented with the sentences containing the vowels in random order. The vowels that were analysed in this study were those found in the following words: bid, bead, beg, bag, bug, bard, pod, board, put, boot and bird. The recordings were carried out in a quiet room, using the Kay Elemetrics Computerized Speech Lab (CSL) Model 4500 at a sampling rate of 22,050 Hz. There were 517 tokens altogether, or 47 tokens per vowel.

Analysis

PRAAT version 4.4.20 (Boersma and Weenik 2006), which is a waveform editor with some speech synthesis and analysis functionality, was used to listen to the sound files, and to view the waveforms and spectrograms. PRAAT was also used to make an orthographic
transcription (see Figure 1), and to measure vowel formant frequencies and vowel durations. In this way, sound, waveforms, spectrograms and transcriptions could all be accessed at the same time.

The formant frequency model was used to analyse the vowels, following the view of Watt and Tillotson (2001: 275) that “the frequencies of F₁ and F₂ [the first and second formants] relative to one another are thought to provide the human speech perception system with the cues necessary for the recognition of individual vowel qualities”. The key speech organ involved in the production of vowels is the tongue, and F₁ and F₂ relate to the degree of raising and retraction of the tongue when individual vowels are produced (Fry 1979: 76–79; Kent and Read 2002: 113). F₁ corresponds inversely to vowel height, while F₂ has a direct correspondence with tongue advancement/retraction (although it is also affected by lip rounding). As illustrated in Figure 2, F₁ increases as vowels are produced lower in the mouth (e.g. compare the F₁ for the vowels in bead and bag) while F₂ increases with the fronting of a vowel (compare the F₂ of the vowels in bead and boot). There are limitations to this formant frequency model but as Watt and Tillotson (2001: 276) point out, “the formant plots can provide an approximate representation of the relative qualities of individual vowels”. This model is commonly used in instrumental analysis of vowels (e.g. Deterding 1997; Watt and Tillotson 2001; Deterding 2003; Hawkins and Midgley 2005).

The F₁ and F₂ of the vowels were measured using linear frequency coding (LPC) tracks, plotted automatically on the spectrogram by PRAAT, after making an auditory examination of the target words and a visual examination of the formant tracks. The F₁ and F₂ of the vowels were measured at the temporal mid-point of the vowel (see Watt and Tillotson 2001: 277; Adank, Smits and Hout 2004: 3; Hawkins and Midgley 2005: 185), which is the point at which a vowel is deemed to be at its most steady state and is least influenced by preceding and following sounds (in this case the stop consonants). The locations at which the measurements were made were recorded to allow for verification. The F₁ and F₂ of
the vowels were measured by two raters and the results of running a Pearson correlational analysis showed a high degree of agreement between the measurements by Rater 1 and Rater 2: $I(515) = 0.99, p < 0.01$ for $F_1$ and $I(515) = 0.99, p < 0.01$ for $F_2$.

Based on average $F_1$ and $F_2$ values for each vowel, $F_1$ vs. $F_2$ vowel plots were generated by converting the values into a Bark scale (Zwicker and Terhardt 1980). The Bark scale was used because it “is thought to be a good approximation of the actual frequency analysis performed by the ear” (Kent and Read 2002: 115). The values were plotted on an $F_1$ vs. $F_2$ rather than a $F_2 - F_1$ vs. $F_1$ plot, although we acknowledge that, as has been known before, there may be an issue regarding the positioning of the back vowels (see Hayward 2000: 159–160). The $F_1$ vs. $F_2$ plot in any case gives a better representation of the traditional vowel quadrilateral than the $F_2 - F_1$ vs. $F_1$ plot.

Apart from $F_1$ and $F_2$ values, the vowel durations were measured (in milliseconds) from the onset to the offset of the vowels based on an auditory and visual examination of the spectrograms. The average values for the measurements were calculated and comparisons were made between vowel pairs to determine if there was a lack of length contrast between the pairs. The results of running a Pearson product moments correlational analysis revealed a statistically significant relationship between the vowel durations measured by Rater 1 and Rater 2: $r(515) = 0.1, p < 0.001$. This means a statistically significant positive relationship existed between these two measurements conducted at different times. Following Deterding (1997), the distance of each vowel from the centroid was determined by the Euclidean distance between the $F_1$ and $F_2$ of each vowel from the centroid which was based on the average $F_1$ and $F_2$ values of all vowels. The average formant frequencies and Euclidean distance (in Bark) and duration (in milliseconds) for the vowels are shown in Table 1.

**RESULTS**

Figure 3 shows the vowel chart for MalE vowels, and it can be seen that there is a lack of contrast between /i/ - /ɪ/, /e/ - /æ/ and /ʌ/ - /ɑ/ while there is more of a contrast between the
Table 1. Average values for F₁ and F₂, Euclidean distance and duration

<table>
<thead>
<tr>
<th>Vowels</th>
<th>F₁ (Hz)</th>
<th>F₂ (Hz)</th>
<th>F₁ (Bark)</th>
<th>F₂ (Bark)</th>
<th>Euclidean distance</th>
<th>Duration (msec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>316</td>
<td>2829</td>
<td>3.07</td>
<td>15.26</td>
<td>4.66</td>
<td>168</td>
</tr>
<tr>
<td>i</td>
<td>373</td>
<td>2653</td>
<td>3.60</td>
<td>14.87</td>
<td>4.04</td>
<td>100</td>
</tr>
<tr>
<td>e</td>
<td>824</td>
<td>2112</td>
<td>7.31</td>
<td>13.45</td>
<td>2.44</td>
<td>112</td>
</tr>
<tr>
<td>æ</td>
<td>895</td>
<td>2078</td>
<td>7.81</td>
<td>13.35</td>
<td>2.70</td>
<td>132</td>
</tr>
<tr>
<td>æ</td>
<td>870</td>
<td>1419</td>
<td>7.64</td>
<td>10.83</td>
<td>1.93</td>
<td>119</td>
</tr>
<tr>
<td>æ</td>
<td>897</td>
<td>1357</td>
<td>7.83</td>
<td>10.53</td>
<td>2.22</td>
<td>211</td>
</tr>
<tr>
<td>æ</td>
<td>806</td>
<td>1200</td>
<td>7.19</td>
<td>9.70</td>
<td>2.26</td>
<td>128</td>
</tr>
<tr>
<td>æ</td>
<td>641</td>
<td>1021</td>
<td>5.92</td>
<td>8.65</td>
<td>2.87</td>
<td>129</td>
</tr>
<tr>
<td>ø</td>
<td>472</td>
<td>1237</td>
<td>4.49</td>
<td>9.90</td>
<td>2.10</td>
<td>82</td>
</tr>
<tr>
<td>ø</td>
<td>410</td>
<td>1026</td>
<td>3.94</td>
<td>8.68</td>
<td>3.41</td>
<td>107</td>
</tr>
<tr>
<td>š</td>
<td>584</td>
<td>1543</td>
<td>5.45</td>
<td>11.39</td>
<td>(0.41)</td>
<td>167</td>
</tr>
<tr>
<td>Average</td>
<td>644</td>
<td>1680</td>
<td>5.84</td>
<td>11.51</td>
<td>2.86</td>
<td>132</td>
</tr>
</tbody>
</table>

Figure 3. Vowel chart of Malaysian English vowels

back vowels, /u/-/ʊ/ and /o/-/ɔ/. This is similar to findings concerning Brunei (Salbrina 2006: 253), and Singapore English (Deterding 2003: 7–8), with the exception of /ʊ/-/ɔ/, which showed less contrast in Brunei and Singapore English. Compared to British English vowels produced in citation form (Deterding 1997: 55), the vowels occupy a smaller vowel space just like Brunei and Singapore English (see Salbrina 2006). Further, the /ɔ:/ vowel in MalE is generally produced lower down than in British English (see Deterding 1997). The average distances from the centroid for MalE, Singapore English (Deterding 2003: 16), Brunei English (Salbrina 2006: 253) and British English (Deterding 1997: 55), vowels were compared. The results suggest that British English (2.90 Bark) vowels are the most peripheral while Brunei English (1.82 Bark) is more central. Based on a correlated samples t-test, the difference between the average distance for MalE (2.86 Bark) and British English is not significant ($t = -0.17, df = 9, p > 0.01$) while it is significant between Singapore...
English (2.41 Bark) and MalE ($t = 3.29$, $df = 9$, $p < 0.01$) and between Brunei English and MalE ($t = 5.243$, $df = 9$, $p < 0.01$).

To obtain a clearer picture of the extent of vowel contrast between vowel pairs in MalE, the distribution of the vowels produced by the individual MalE speakers are shown in Figures 4 to 8. As can be seen, all the vowel pairs display overlapping distributions indicating a lack of real contrast, especially in the case of the vowels of bag and beg, and bud and bard, which are essentially produced as the same vowel. Correlated samples $t$-tests indicate there is a significant difference in the average $F_1$ and $F_2$ between /i/ and /iː/ ($t(46) = 5.36$, $p < 0.01$; $t(46) = 5.99$, $p < 0.01$); /o/ and /oː/, ($t(46) = 11.54$, $p < 0.01$).

© 2010 The Author(s). Journal compilation © 2010 Blackwell Publishing Ltd.
p < 0.01; $t(46) = 11.16, p < 0.01$); and /ʌ/ and /ʊ/ ($t(46) = 5.5, p < 0.01$; $t(46) = 10.23, p < 0.01$). For /e/ and /æ/, there was a significant difference in the $F_1$ ($t(46) = 6.19, p < 0.01$) but not the $F_2$ ($t(46) = 1.64, p > 0.01$) while there was no significant difference for $F_1$ ($t(46) = 1.91, p > 0.01$) but with a significant difference in $F_2$ ($t(46) = 3.47, p < 0.01$) between /ʌ/ and /ʊ/.

A comparison among the three ethnic groups suggests that while all the groups display the typical lack of contrast between the traditional vowel pairs /i/ - /ɪ/, /e/ - /æ/, and /ʌ/ - /ʊ/, there appear to be differences in the extent of the contrast. For example, a comparison
of Figures 9 to 11 shows that the Malay speakers have the least contrast between /u/ and /iː/, while the /æ/ and /æː/ vowels appear to have merged among the speakers of Indian origin. This could be due to the speakers’ first language but more detailed analysis needs to be done to prove or disprove this. The average distance from the centroid for the vowels produced by each ethnic group was calculated as Malay (2.91 Bark), Chinese (2.82 Bark) and Indian (3.00 Bark), suggesting that the latter was more peripheral than the others. Correlated samples t-tests show no significant differences between the average Euclidean distances from the centroid for the three ethnic groups: Malay and Chinese ($t(9) = 0.81$, ...
The lack of contrast in terms of vowel quality confirms the current descriptions of MalE vowels. However, contrary to current descriptions of MalE, the differences between the mean durations of each of the vowel pairs were all statistically significant (/i/ and /i:/: \( t(46) = 11.05, p < 0.01 \); /e/ and /æ/: \( t(46) = 7.12, p < 0.01 \); /ʌ/ and /œ/: \( t(46) = 15.92, p < 0.01 \); Malay and Indians \( (t(9) = 1.34, p > 0.01 \); Chinese and Indians \( t(9) = 2.25, p > 0.01 \)).
/u/ and /u/; t(46) = 7.79, p < 0.01) with the exception of /u/ and /ɔ/, t(46) = .34, p > 0.01, which incidentally showed more contrast in terms of quality.

DISCUSSION

One of the interesting observations to be made is that the contrast between vowels in Malaysian English does not match the classical notion of phonemic contrast. The overlap in the vowel plots do not indicate an absolute phonetic difference, but rather an overlap area of phonetic similarity. Within this overlap area, we cannot assume that no distinction is made, because other phonetic features could be involved. In most cases, there is a significant difference in duration. In the case of pod and board, instances of the vowel in pod appear to stray into the area normally used for the vowel of board. If we consider the speech of the undergraduate subjects as a whole, there would appear to be a partial merger between the two vowels. At the individual level, on the other hand, it would still be possible for the contrast to be maintained consistently. A very similar situation for these two vowels is also found in some varieties of Scottish and Irish English, and in American English.

In this paper, we have analysed acoustically the vowels of 47 female undergraduates as an undifferentiated group. All informants appear to have acquired vowels with new qualities, namely, qualities that do not match vowels in Malay, Chinese or Tamil; the main L1s for the group. The figures for the group indicate overall that these speakers make length contrasts in the vowels, even though this is not a property of vowels in Malay or the main varieties of Chinese. In other words, these speakers would appear not to be recycling mother tongue vowels for use in English – which is what one might expect in the case of less proficient language learners – but have learnt to produce distinctively English vowel sounds. These conclusions apply of course only to the group of undergraduates who happen all to be female: we cannot draw any conclusions at this stage about the speech of male students.

This study is concerned only with the vowel system itself. However, it is well known that differences in vowel duration are used not only to distinguish vowels, but also voiced and voiceless consonants as in beat and bead. The fact that our informants use duration to distinguish vowels does not of itself mean that they are able to manipulate duration as a means of establishing the voicing contrast. This remains to be investigated in future research.

CONCLUSION

English vowels are conventionally classified according to phonetic quality and duration. We have found that the mean formant values for our informants differ slightly from those reported from Singapore, Brunei and RP. This is of course to be expected, and it would be quite extraordinary if it were not so. We would also expect slight differences in mean formant values if we were to take samples of speech, for example, from different towns in England. We have also found that with the exception of the vowels of pod and board, vowel quality distinctions are supported by differences in duration. Again, we can expect slight differences in the durations used by our informants and those in use in other varieties of English. The main point, however, is that the same two parameters are used to maintain phonemic contrasts in the vowel system. The slight differences in the values of these two parameters reflect trivial differences in what is essentially the same system.
Further research is required to ascertain whether or not our findings can be generalized to parts of the vowel system (such as diphthongs), which we have not yet investigated. However, according to our findings so far, it would appear that the vowel systems used by our informants do not differ significantly from those in related varieties, such as Singapore and Brunei English, and RP. One of the goals of our research is to find ways in which the spoken English of our informants differs significantly from comparator varieties, but we are able tentatively to exclude the vowels, or at least the monophthongs, of the English phoneme system.

REFERENCES


(Received 8 January 2009.)