The Spoken English of Hong Kong: A Study of Co-occurring Segmental Errors

Richard Stibbard
School of Arts, University of Surrey, Guildford, UK

There is broad agreement as to many of the segmental features of the Hong Kong accent of English: neutralisation of vowels which contrast in Standard Southern British English or General American, non-release of final stops, simplification of consonant clusters and devoicing of coda consonants. However, while it is apparent that there is no reason why these features should not co-occur within single words, such co-occurrences have not been identified in previous studies, perhaps because treatments of HK pronunciation have generally used lists of words and have thus elicited atypically careful pronunciation. The connected speech data used in the present study indicates that findings from word lists may not apply to more naturalistic speech. In this study, speakers produced many words with more than one segment sounding like another English phoneme, sometimes affecting all the segments of a word. Although overt signs of misunderstanding hardly arose, this indicates merely that the lack of such overt signals is no sign of acceptability. Arguments that Hong Kong English pronunciation should be viewed as 'phonological' in its own right are rejected as inappropriate, both on grounds that this interpretation is not supported by the phonetics of the data, and more conclusively on sociolinguistic grounds.

Keywords: Hong Kong English, pronunciation, phonology, segmental errors, phonemic overlap, intelligibility

The Hong Kong accent of English has been described by a number of previous writers, including Luke and Richards (1982), Bolton and Kwok (1990), Chan and Li (2000), Hung (2000) (reprinted as Hung, 2002), and Peng and Setter (2000). Many of the features which these writers describe are widely agreed upon, and most are clearly due to interference from Cantonese.

Vowels

Features of Hong Kong English which are widely reported include a reduced pure vowel inventory compared with that of native speakers. Often the length and quality contrast between the long and short, or tense and lax, vowel pairs which are distinctive in Standard Southern British English (SSBE) and many other native accents such as General American (GA) are not realised reliably. Because Cantonese has no such distinctive pairs, these are typically neutralised such that, for example, the distinction between 'beat' (/biːt/) and 'bit' (/bɪt/) is not reliably made. The same applies to the pairs /ɒ/ *vs.* /ɔː/ and pairs /ʊ/, *vs* /uː/ and to the quality difference between /e/ and /æ/ as in 'met' *vs* 'mat'.

There is, however, an important difference of opinion as to exactly what Hong Kong speakers do produce in cases such as these. Hung (2002) believes that in these cases both are consistently pronounced alike, e.g. both 'bit' and 'beat' as [bit], and uses this claim to support his argument

Vol. 17, No. 2, 2004

that Hongkongers' realisations should be accorded phonemic status. Hung's claims of homogeneity among his subjects in the failure to produce a formant/duration distinction between the long/short phoneme pairs of SSBE/GA are supported by standard deviations showing inter-speaker variability, but in his subsequent discussion he discounts this and averages across speakers. Chan and Li (2000), in contrast, report that their subjects' realisations of these phonemes varied between the two native speaker targets, sounding like one or the other, but not necessarily the right one, or a sound intermediate between the two; the present study lends weight to Chan and Li's (2000) view of instability rather than stability.

Less well agreed is how Hongkongers produce English diphthongs. With the exception of a curious claim that Hongkongers make a *phonemic* distinction between /aɪ/ and /ʌɪ/ in words with SSBE and GA /aɪ/, Hung (2002: 129) reports that contrasts are maintained as in SSBE. Chan and Li (2000), in contrast, report widespread monophthongisation, [ɒ] for /ɔɪ/, [e] or [æ] for /eə/, and [e] for /eɪ/, but also the pronunciation of diphthongs as separate vowels with an intervening glottal stop: [pʊʔa] for 'poor'. Bolton and Kwok (1990) report /eɪ/ realised as [ɪ]. Again this indicates instability in the Hong Kong accent rather than a stable system. The different findings may also be due in part to the different methods of data collection and analysis used. Hung's use of word lists read aloud may have led to unusually careful and accurate pronunciation and his practice of discounting inter-speaker variability may obscure individuals' pronunciation problems. Luke and Richards and Bolton and Kwok used connected speech data, which may have avoided this problem and given more representative results. Chan and Li's paper is based on informal observations in the classroom of connected speech but can be criticised for the lack of careful repeated listening to audio data and the possibility that remembered observations may be inaccurate.

Consonants

As Cantonese has a smaller inventory of consonant contrasts than English, contrasts made by native speakers of English are often lost as sounds are substituted from the Cantonese phoneme inventory: [f] for /θ/, [d] for /ð/, [w] or [f] for /v/, and [s] for /ʃ/. Because word-initial [l] and [n] are often in free variation in Cantonese, as in 你 ('you') pronounced [neɪ] or [leɪ], the two English sounds are often pronounced in free variation. Chan and Li report that realisations of English /r/ and /w/ also enter into this free variation, with word-initial /r/ pronounced as [l] by some speakers and as [w] by others. In word-final position, dark 'l' is often replaced by [u] and /n/ is often not pronounced.

Making the distinction between voiced and voiceless consonants in the coda is another well-known difficulty, especially in word-final position, as the three Cantonese word-final stops, /p/, /t/, and /k/ are [p˺], [t˺], and [k˺], i.e. voiceless and unreleased. Often, these or a glottal stop are substituted for both their English voiced and voiceless counterparts. All other voiced consonants are also problematic: devoicing is widely reported in the previous studies.

Another problematic feature of English for Cantonese speakers is its complicated syllable structure compared with the simple Cantonese syllable

structure, (C)V(C). This typically results in substitution or deletion of consonants in clusters. Peng and Setter's (2000) paper is an investigation of the deletion and retention of consonants in clusters; the authors conclude that this has a regular pattern according to the morphosyntactic structure. However, they also point out that if two different target words are made to sound the same, this can lead to serious intelligibility problems, which they believe should be a priority for correction in the classroom (2000: 106).

Bolton and Kwok (1990) report that while Hongkongers' vowels are generally closer to those of SSBE than to General American, some features of American English may be heard from some speakers, such as flapped 't' in words such as 'city' and the American vowels in 'job', and 'dance'.

The Study

The purpose of this study is to investigate the co-occurrence within single words of errors involving gross phonemic overlap, as well as place of articulation shifts and insertion of intrusive consonants. Phonemic overlap has been widely reported in previous studies, but has previously only been identified as occurring at the rate of one per word. The present data indicate that the problem is made much more serious than previously thought by the fact that such errors can co-occur. The decision to concentrate on these errors and their co-occurrences was made on the basis of frequency of occurrence in these data and because of their evident negative effect on the intelligibility and acceptability of the Hong Kong accent.

The data used are a collection of audio recordings made in 1997 of the speech of undergraduates studying on BA and BSc courses at Hong Kong Baptist University. Seventeen students volunteered to take part in the study, four male, 13 female. All had taken the HK Examinations Authority Advanced Supplementary Level in Use of English, obtaining the following overall grades: B (one student); C (three students); D (ten students); E (three students).

The data collection method elicited speech partly in interaction, partly monologue. The rationale behind this method was to record for analysis more naturalistic connected speech than was used in previous studies while retaining some control over the words used and style of speech. The subjects took part in three spoken activities: two information-exchange activities (interaction), and retelling a short story from memory (monologue). The information-exchange activities were a map-reading task and a pegboard description task. In the map-reading task a pair of subjects had maps representing the same place but with different amounts of information available. One of the subjects, the giver of information, had a route marked on it to describe to the other, who had to draw the route on his/her map, asking questions as necessary so as to clarify the route.

In the pegboard description task, the roles were reversed, the information-giver from the map-reading task becoming the receiver of information. The subjects were supplied with a pegboard on which the information-giver had a number of coloured pegs around which were stretched coloured elastic bands forming geometric shapes. The task was to exchange information so that the receiver could reconstruct the shapes and colours as they were on the other

subject's board. The subjects were seated such that they could see each other but not the other's map or pegboard respectively.

For the story-retelling task, a short written story was provided. The subjects were given time before recording began to read this and remember the main points. Later they were asked to retell the story from memory in a spoken style. Acoustically high quality recordings were ensured by recording in a sound-proofed recording studio direct-to-disk and simultaneously onto DAT. In total, 9606 seconds of data were analysed.

Analysis

Perceptual phonetic analysis was carried out by the author by repeated listening to the audio recordings using the phonetic analysis software Praat (Boersma & Weenink, 2003) and its time-aligned transcription facilities. Initially, transcriptions were made of occurrences where gross errors involving phonemic overlap occurred. For this paper, only those cases were taken into account where several cases of phonemic overlap co-occurred within a single word, or intrusive consonant sounds were inserted, or place of articulation shifts occurred. The latter is a case of phonemic overlap where co-occurrence is not always involved, but was included because of the gross perceptual effect of the resulting errors.

The use of perceptual analysis is felt to be justified because an instrumental approach is largely unhelpful when naturalistic data are used. It is possible to measure variables such as vowel duration and formant values accurately only when pronunciation is tested artificially in carrier words designed to maximise the clarity of the consonant/vowel/consonant divisions, which rules out collecting data representative of connected speech in a communicative setting. Even when instruments are used to support and document findings, the perceptual effect of the speech should remain the deciding factor.

Although convincing arguments have been made that native speakers should not always be the arbiters of correctness, particularly where no native speakers are involved in a conversation (Jenkins, 2000), it is felt that this approach is justified by the situation in Hong Kong, where communicating with native speakers, mainly in a professional context, is a very important use of English.

Findings

The present data generally support the findings of the studies reviewed above: features such as the failure to distinguish between long and short vowel phonemes, substitution of sounds from Cantonese, non-release of final plosives and devoicing of voiced consonants are widespread. They also support Chan and Li's view of instability in vowels rather than Hung's claim of stability. For instance, vowels which contrast in length or tenseness in SSBE and GA are not consistently produced as a single intermediate vowel as Hung reports, but are on occasion pronounced very long and tense, on others short and lax, not necessarily correctly for the intended word, and on other occasions intermediate between the two.

Co-occurrences of phonemic overlap

Where the present findings differ most from those of other writers on Hong Kong English is that there seems to be no potential limit to the number of co-occurrences of phonemic overlap in a single word, up to and including all the phonemes of a word. Perhaps because of the preference for word lists, writers such as Hung (2002) and Peng and Setter (2000) may have elicited atypically careful speech in which this natural tendency in informal speech was not present.

Brown (1995) reports similar effects in Singapore English and gives high priority to it in teaching because of the very severe effects on intelligibility. His subjects produced utterances such as 'There's a butterfly on your [pe?]', which in the context, a camping trip, could have referred to a bag, a back, a pack, or a peg and was in effect unintelligible.

In the present data there is in addition a large amount of apparently unsystematic variation as to what features appear in the speech of particular speakers and within the same speaker's speech. Many of the realisations are extremely distant from what a native listener would expect to hear and can only be identified because of the context. Even when they can be identified from the context, they are often indubitably wrong and would not be accepted in any variety of English.

There are 199 occurrences in the data of words affected in a similar way in the speech of more than one speaker. Examples are found in the speech of all the speakers except two. The words most often affected are not uncommon words which might be unknown to the speakers: they are all common words, such as *walk*, *bridge*, *pond*, *board*, *black*, *blue*, *north* and *straight*. An example is found in the various pronunciations of the word 'bridge' (SSBE and GA /brɪdʒ/). In all, eight out of the 17 subjects pronounce 'bridge' with more than one phoneme incorrect.

Taking the segments of the word 'bridge' in order, the first to cause widespread problems is the /r/. Previous writers would not lead one to expect deletion of the /r/ but rather substitution by a sound closer to English /l/, especially in clusters. This is produced by one speaker, but deletion is the commonest pronunciation. This is often combined with inconsistent vowel length and quality plus devoicing of the final consonant. All these features have been previously identified; but never in combination. The result is that [bitʃ] and [biːtʃ] (sounding like SSBE 'bitch' or 'beach') are both produced on multiple occasions. Also found is [bɪtʃ], sounding exactly like SSBE 'bitch'. Idiosyncratically, the word is once pronounced as [pʰiːtʃ], with a clearly aspirated initial plosive, no [r], perfect realisation of SSBE /iː/, and final devoicing. The result sounds exactly like SSBE 'peach'. The following mispronunciations of the phonemes are co-occurring within one word:

/b/ as [b] or [pʰ]
/r/ as [r] or [l] or Ø
/ɪ/ as [ɪ] or [i] or [iː]
/dʒ/ as [dʒ] or [tʃ]

Although some of these occurrences, such as [pʰiːtʃ], are idiosyncratic, co-occurrence of deviations from SSBE/GA involving overlap onto another SSBE/GA phoneme is common. It seems as though there is no limit to the

extent to which these deviations may co-occur, up to and including all the seg-ments of the word. The fact that some of the realisations are idiosyncracies personal to certain subjects should strengthen, rather than lessen, concerns.

Another example is found in the various realisations of the word 'width' (SSBE and GA /wɪdθ/) found in the present data. Previous studies lead one to expect difficulties with this word, but do not predict problems on the scale encountered here. Chan and Li's (2000) contrastive analysis indicates that Hongkongers may not distinguish /ɪ/ as in 'width' from /iː/ as in 'weed', that /θ/ is a difficult sound which may be realised as [θ] or [t] or [f], and that consonant clusters in general present problems, commonly resulting in deletion of one of the consonants. Chan and Li report difficulties with English /r/, this being realised as [w], but not the reverse.

However, in the present data, all these possible realisations combine in one word, inconsistently from speaker to speaker and from utterance to utterance. Realisations of the word include: [wɪdθ], [widθ], [wiːdθ], [wiːf], and [rɪf].

This is a large number of possible realisations for the listener to cope with. It is interesting to speculate how many realisations are possible if all these mis-pronunciations could co-occur without constraint. The relevant formula is:

$$n = c_1 \cdot c_2 \cdot c_3$$

where n is the number of realisations of the word, and c is the number of rea-lisations of each segment, repeated for the number of segments in the word, the subscript indicating the position of the segment in the word.

Thus, in the case of the word 'width' above, where /w/ was realised as either [w] or [r] ($c_1 = 2$), the vowel as [ɪ], [i], or [iː] ($c_2 = 3$), the /d/ as /d/ or Ø ($c_3 = 2$), and the /θ/ as [θ] or [f] ($c_4 = 2$), the maximum potential number of realisations is $2 \cdot 3 \cdot 2 \cdot 2 = 24$:

[wɪdθ]	[wɪdf]	[wɪθ]	[wɪf]
[rɪdθ]	[rɪdf]	[rɪθ]	[rɪf]
[widθ]	[widf]	[wiθ]	[wif]
[ridθ]	[ridf]	[riθ]	[rif]
[wiːdθ]	[wiːdf]	[wiːθ]	[wiːf]
[riːdθ]	[riːdf]	[riːθ]	[riːf]

Even with the admission that this is an artificial mathematical exercise designed to draw attention to the scale of the problem, and that many of these realisations do not occur in the present data, it seems that there is no system-atic constraint on them. It is surprising if one reads through the list just how many of them will sound authentic to those familiar with the speech of Hong Kong learners. Another example of such errors co-occurring in one word is the pronunciation of 'walk', which is realised variously as [wɔːk˺], [wɔk˺], [wɔkʰ], [wɒkʰ], [rɔkʰ], and [rɒk˺].

The realisation of 'walk' with an initial [r] such that it sounds like 'rock' occurs in the speech of five of the 17 subjects. Only once is it corrected, when a subject repeats the word, correcting the initial [r] to a [w], and inserting an unreleased final stop where there was none, but not correcting the vowel, so that the result sounds like 'wok', which is not perfect but, given the

context, is much clearer: 'do I walk through ([rɒfruː]) … WALK through' (['wɒkˀ fruː]).

The word 'walk' occurs most often during the map-task because one subject had to guide the other to 'walk around the Lion Rock' as marked on the map. A combination of errors produced four times by one speaker and echoed once by her interlocutor was the pronunciation of 'walk' with [r] and 'lion' with no initial /l/ as [aɪən], giving the effect of 'rock around the iron rock'.

Contrastive analysis leads us to expect that in 'north' the initial /n/ might be pronounced as either [n] or [l], that the vowel might variously be pronounced close to that in 'not' [ɒ], or more like an SSBE speaker's 'nought', as [ɔː], or intermediate between the two, as [ɔ], and that the /θ/ might be realised as either [θ] or [f]. But it does not lead us to expect that all of these will occur, yet this is what happens in the present data, where [nɔːθ], [nɒθ], [lɔːθ], and [lɔf] all occur, as well as [lɔːθtʰ] and [nɔːθtʰ], idiosyncratically in the speech of one subject.

The formula for the maximum number of possible realisations excluding the inserted [tʰ] shows that there are $2 \cdot 3 \cdot 2 = 12$ possibilities:

[nɔːθ]	[lɔːθ]	[nɔːf]	[lɔːf]
[nɔθ]	[lɔθ]	[nɔf]	[lɔf]
[nɒθ]	[lɒθ]	[nɒf]	[lɒf]

Again, not all these possibilities occur in the present data, but enough do to show again that more than one such error can occur together and that some of the combinations result in realisations which are phonetically distant from native listeners' expectations, although none in this case sound like another word.

Combination of errors do not always occur at the word level: they may spread across phrases. An example occurs in one subject's attempts to say 'bamboo pipe', realised as [bæŋbuːn pai] and [bækbuːŋ pai], both of which would probably defy understanding even in context. In two other subjects' speech, 'pipe' is realised as [taitʰ] and [paikʰ], while the call from two subjects to help the '[íliɗi guːps] (needy groups) in our society' might also present a challenge to the listener.

As the meanings (to the ears of native listeners at least) of the resulting realisations are a matter of chance, it is not surprising that at times unfortunate results can occur. An example which could have been embarrassing or worse had it occurred in a situation where the stakes were higher, such as a professional presentation, is in the utterance 'there are many ['pinəs] come from China' ([p] indicates a voiceless unaspirated stop). The target word was 'business' but with [i] substituted for /ɪ/, a deleted /z/ in the first syllable, and the plural /-ɪz/ missing, the resulting word sounded entirely unlike what was intended. A similarly embarrassing result occurs in the utterance by another speaker 'what is the meaning of the blue ['pʰiːnis] in the middle'. The target utterance was 'pin is', ('what is the meaning of ''the blue pin is in the middle''?'. The unusual use of a direct question embedded in another question, combined with devoicing of the final [-z] and the long [iː] for short /ɪ/ and a failure to distinguish stressed and unstressed syllables combined to make the pronunciation perfectly clear but not what was intended.

Where it seemed that mispronunciations were due to lexical unfamiliarity, they were excluded from the analysis. An example is *giraffe*, realised variously

as [ʤɪˈræf], [ʤɪrɪf], [ʤrɪf], [gɪrɪf], [grɪf], and [ˈgrafi]. The repeated realisation of *penguins* as [ˈpɪʤɪnz] was excluded on the grounds that it was a lexical error not a pronunciation error, as was the very common and well-known error, [bɪə] for *bear*, because of the special difficulty the spelling presents. It is never-theless saddening that subjects who have studied English throughout the Hong Kong school system to university level still do not know the pronunci-ation of this rather elementary word.

On occasions these phonemic overlaps take the form of place of articulation shifts involving word-final stop consonants. There are examples of shifts from alveolar to velar (e.g. *straight* [streɪkʰ]), and velar to alveolar, (e.g. *like* [laɪtʰ]), from bilabial to alveolar (e.g. *pipe* [paɪtʰ]), and from bilabial to velar (e.g. *pipe* [paɪkʰ]). Thus all possible shifts occur except shifts to bilabial. Again, all occur on common, mainly monosyllabic words such as *board* [bɔkᵇ], *put* [pʊkʰ], *red* [rɛkʰ], *map* [mætʰ], *pipe* [paɪtʰ], *back* [bætʰ], and, combined with deletion of /l/, *black* [bætʰ].

It is apparent that these shifts have the effect again of making a word sound quite unlike the target. This is the more so because the pronunciation of the incorrect final stop is often strikingly precisely articulated, released and aspirated, making it sound entirely unlike a foreign accent, but like a perfectly pronounced English word, sometimes the wrong one, sometimes a nonsense word which would be permissible according to the phonotactic rules of English.

An example of this, which also happens to be one of the very rare instances of self-correction, is where a speaker corrects what sounds exactly like SSBE or GA 'the stucking point' [ˈstʌkʰɪŋ] to 'the starting point' [ˈstɑtʰɪŋ].

Another prominent feature of the data is the repeated insertion of consonant sounds [tʰ], [s] and [kʰ], as in 'however' pronounced [haʊˈevə tʰ]. This is an extremely common occurrence: there are 197 such inserted consonant sounds, in the speech of all but four of the speakers. The commonest such sound is [tʰ], which occurs 158 times, in the speech of 13 out of the 17 speakers. Less com-mon is [s], which occurs 34 times, produced by eight speakers, and [kʰ], which occurs five times, produced by three speakers.

There is some phonetic patterning to these insertions. The two alveolar insertions occur most commonly after an alveolar consonant: 75 out of 158 [tʰ] insertions (47%) and 20 out of 34 [s] insertions (59%). Most common of these preceding consonants is [n], which accounts for all 20 of the post-homorganic [s] insertions and 58 out of 75 of the post-homorganic [tʰ] inser-tions. In the case of the [kʰ] insertions, there are too few of them in the data for any patterns to emerge.

The commonest following environment is a pause, accounting for 101 of the 158 [tʰ] insertions (64%) and again 20 of the 34 [s] insertions (59%). The following are thus typical examples (the symbol | indicates a tone unit boundary).

| and my option [tʰ] is |
| it is in a straight ([strek]) line [tʰ] |
| the red pin [tʰ] on the left |
| then [s] we make the third one |

| gate one [tʰ] |
| all the animals [tʰ] |

The second commonest preceding environment is a vowel (58/158 or 37% of the [tʰ] insertions and 8/34 or 27% of the [s] insertions). Examples of these are:

| however [tʰ] |
| finally [tʰ] |
| so [tʰ] |
| one day [tʰ] |
| actually [tʰ] |

These often coincide with a tone unit boundary following the inserted sound. In many cases these inserted sounds do not result in one word sounding like another and so probably do not contribute directly to intelligibility problems. However, this is dependent on the lexical items in the environment: on several occasions the word 'can' precedes a [tʰ], potentially resulting in confusion with 'can't'. In these data, the vowel targets for 'can' and 'can't' are obviously the SSBE ones, [kæn] and [kɑːnt] respectively, so intelligibility is not affected. Sometimes Hong Kong learners aim for the most part at SSBE vowels but substitute General American targets for certain words, such as 'dance' [dæns] and 'can't' [kænt]. This inconsistency could cause intelligibility problems if combined with the inserted [tʰ]. Whether or not intelligibility problems are caused, the sounds are certainly unusual, disrupt fluency and are a distraction for listeners.

Discussion

The use of relatively natural, connected speech in preference to word lists entails a loss of control over the phonetic environments of the data, meaning that accurate instrumental analysis of durations and formants is problematic. The analysis used here could be criticised for picking a series of errors from out of the speech in a hypercritical way. However, these errors are quite gross, involving clear phonemic overlap, not small details of accent, which make words sound entirely unlike the target.

Shortcomings of a phonological treatment of the Hong Kong accent

Hung's claim (2002) for a phonology of Hong Kong English is based on Mohanan's (1992) treatment of Singapore English. He follows Mohanan's view that 'the prescriptive approach is unhelpful, and that it runs the risk of distorting the true nature of the system' (Hung, 2002: 119). However, writing for the Hong Kong audience, which might be less receptive to claims that 'linguistic emancipation' are high on the agenda, Hung judiciously omits the more overtly political language of Mohanan, who terms the contrastive approach 'parasitic' (1992: 111), and claims that pronunciation is a way of asserting ideological freedom from 'the cultural and intellectual bonds of colonialism' and 'a symbolic rejection of British colonialism' (1992: 112).

Hung's argument rests on the evidence he presents that realisations are stable. In the case of the 'long/short' vowel pairs (/i:/ *vs* /ɪ/, etc.), this means that a single intermediate realisation is always produced, as shown in Figure 1.
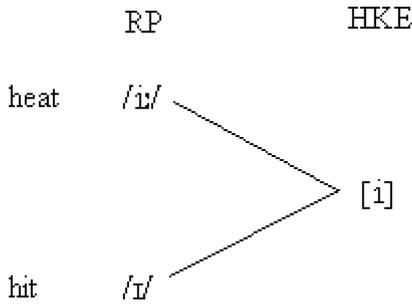
RP   HKE

heat   /iː/

                                          [i]

hit   /ɪ/

Figure 1 HK English vowel realisations as a consistent sound intermediate between two SSBE or GA phonemes

RP   HKE

heat   /iː/   [ɨ]

                            [i]

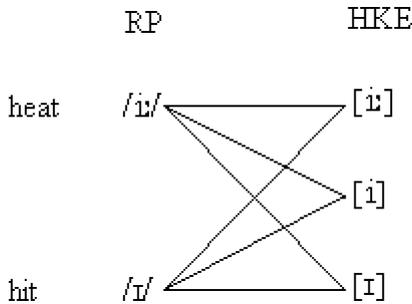hit   /ɪ/   [ɪ]

Figure 2 HK English vowel realisations as inconsistent sounds varying between SSBE, GA vowel sounds

In contrast, Chan and Li (2000) state explicitly that this is not the case in their data, and that vowel realisations are inconsistent, varying between either extreme of the SSBE/GA members, not necessarily the right one, or an intermediate sound, as in Figure 2.

If the case depicted in Figure 2 is a more accurate representation of what happens in connected speech, then the argument that Hongkongers' vowel realisations should be seen as phonemes is difficult to sustain, as it appears that they are simply aiming inconsistently and with varying success at native speaker models. The variability found in the present analysis and the descriptions given by Chan and Li (2000) indicate that this is a more accurate description of the reality of connected speech.

Turning to consonants, Hung again claims that in most cases Hong Kong realisations are stable, but the claim encounters more obvious problems because of the discrete nature of consonants rather than the scalar nature of vowels. Trying to explain the realisation of /v/ as [w] in syllable-initial position and [f] elsewhere, for example, he rejects the intuitive explanation that the variation is due to different strategies for coping with a difficult sound in different environments, and attempts to explain both as phonemes, thus producing a description unlike the phonology of any other accent of English.

Hung raises hopes that his approach will be revealing of the learner's mental processes and criticises the comparative approach for being psycho-linguistically unrevealing:

> ...upon noticing that many HK learners pronounce the word *net* as [let], or *let* as [net], teachers often say that they 'confuse [l] and [n]. To say that learners confuse two categories (either in phonology or syntax) is not very illuminating. Do they have two mental representations, or one? (Hung, 2000: 337)

But the promise of psycholinguistic enlightenment is not fulfilled. As others have before him (Chan & Li, 2000), he finds [n] and [l] in free variation for English /n/ and /l/: both occur, but not necessarily in the right place, and no further explanation is offered.

Hung's methodology is self-fulfilling because in the process of assigning phonemic status to sounds, he does away with the need for detailed phonetic listening: the abstraction /i/ subsumes the continuum of sounds [iː – i – ɪ – ɪː]. To work in this way, to hear the language through a preconceived phonological filter with phonetic details ignored, will necessarily support the preconception without testing it in any meaningful way. Only by approaching the data without such theoretical preconceptions can the true nature of the accent be gauged. The phonetic facts, with the details of individuals' speech exposed, should be established before abstractions are posited.

Possible explanations of the errors

A possible explanation for these errors is that they are simply idiosyncratic, akin to slips of the tongue. However, this is an inadequate explanation in view of the qualitative differences between slips of the tongue and the processes in these data. Slips of the tongue are well described, and are transient, usually corrected immediately and involve processes such as misordering, omission and replacement of units in a string (Boomer & Laver, 1968). An example is: 'Our frunds . . . funds have been frozen', where the misordering is at the segmental level and the realisation of 'funds' as 'frunds' is due to the influence of the /fr-/ of 'frozen' later in the utterance.

Brown (2000) reports on slips of the tongue in Singapore English, and finds that they are qualitatively the same as those made by native English speakers. As in Boomer and Laver's data, they involve misordering, replacement, or omission of segments or sequences and are usually corrected immediately. Examples are:

> 'He was hanged for trug . . . drug traficking'
> 'You are supposed to kish . . . kiss the fish'
> 'Let's go flight-kying . . . kite-flying'
> 'Safety clin, paper pip . . . safety pin, paper clip' (Brown, 2000: 32)

The features in the present data involve processes quite unlike these, being due not to misordering of correct phonemes but to repeated incorrect pronunciation of phonemes, and are almost never corrected.

A simpler and more intuitive explanation is the traditional one that the attempt is being made to speak English with the segmental inventory of

Cantonese. As English relies on a large number of phonemes for its semantic contrasts, this results in multiple homophony. Increased homophony means less intelligibility, which demands more indulgence from the listener. This may or may not be forthcoming.

Intelligibility and the listener

It might be expected that gross mispronunciations of the sort documented would give rise to self-corrections, corrections or requests for clarification by the listener, or overt breakdowns in communication. In fact, there are very few instances in the data of any of these and conversations continue without comment despite words and phrases being obviously mispronounced.

A likely explanation for this is that the type of activity used to elicit the speech, information-transfer 'communicative activities', is very constraining of lexical choices at any point in the conversation. For example, as the map contained no bitches, beaches, or peaches, but had two bridges, extreme mispronunciations of the word 'bridge' could still be understood by the listener because the closest approximation to them on the map was 'bridge'. Likewise woks, rocks, irons and lions. In Brazil's (1997) words, the existential paradigm was so tightly constrained that it was in effect a paradigm of one, thus allowing grossly inaccurate pronunciation to pass.

If this explanation is correct, then the value of this type of classroom 'communicative activity' is in doubt, as the constraints of the situation allow speakers to muddle through with much less accurate pronunciation than is possible in the wider existential paradigm of the real world.

Other factors, such as a widespread unwillingness in most cultures to correct another speaker overtly mean that such corrections probably take place rather rarely. Even when another person's speech is not clear, polite listeners often prefer to pass over difficulties in the hope that misunderstandings will be temporary and that the context will clarify them. Often it does, but when it does not the opportunity to ask for clarification without embarrassment may have passed.

It is thus apparent that, while communication breakdowns are evidence of incorrectness, the reverse does not hold: lack of overt comment by participants is not evidence of satisfactory use of the language, and the question arises as to how these speakers would have fared in conversation with listeners (either native or non-native speakers of English) unfamiliar with Cantonese pronunciation of English. Arguments that listeners can disambiguate faulty pronunciation on the basis of context (top-down processing) are no doubt true some of the time and in some situations, but this amounts to guessing and it depends on the good-will of the listener as to how much mispronunciation will be tolerated. The extent to which good-will is extended depends on many factors such as participants' cultures and relative power, status and ages. If the young Hong Kong graduate with a rather shaky command of English phoneme contrasts is in a high-stakes, high stress situation such as a job interview, he or she cannot afford to rely on good-will, but should have been prepared to speak to an internationally intelligible standard. Relying on guesswork from the listener is in any case inconsiderate and weakens the speaker's position

and professional image, even if it is possible with imagination and tolerance to guess what is meant.

It is also apparent that non-native speakers are likely to face as much or more difficulty than native speakers of English in understanding faulty pronunciation, because their perception is filtered through the transfer effects of their own first language. A good example of the confounding effects of transfer from two first languages occurs in Jenkins (2002: 90). A Japanese speaker attempting to say 'red cars' is heard by a German listener as saying 'let cars', who interprets this as 'hire cars' and signals incomprehension. The reasons are clear: the Japanese cannot produce the distinction between English /l/ *vs* /r/, while the German cannot perceive final voiced from voiceless stops, a contrast which does not occur in German, even though the speaker pronounced the final consonant correctly. Neither was aware of the pronunciation and perception problems caused by the other's transfer effects, and communication broke down. A more amusing example given by Jenkins (2002: 88) is 'Shakespeare's bathplace' (for 'birthplace') produced by a Japanese speaker in a talk and not understood by anyone except by other Japanese. In this case incomprehension was not signalled and the student's talk left most of the audience wondering where the bath fitted in.

Jenkins' (2002) data show that the assumption that listeners can use top-down contextual information to disambiguate phonetically unclear utterances is incorrect. She shows that non-native listeners in particular rely heavily on what they hear and are unable to understand even when the context would appear to make only one meaning possible.

Sociolinguistic considerations: Attitudes towards the Hong Kong accent

Phonemic status cannot be assigned solely or even primarily on the basis of the speech signal. Of more importance than phonetic details is the attitudes of speakers to their accent. If speakers feel theirs is a recognised, stable variety which can stand independent of other varieties, this gives the strongest support to the demand for a phonological description.

However, those who, with Hung, propose that we view Hong Kong English as a legitimate 'new variety' (Bolton, 2000, 2002; Bolton & Lim, 2002) appear unaware of the work on attitudes to the accent by Luk (1998) and Tauroza and Luk (1997). In these studies it was found that Hong Kong children rated their own accent lower than RP in all ratings concerned with professionalism, attractiveness and even empathy. Luk (1998) writes that there was no evidence of a desire for 'linguistic emancipation', but that the native model was overwhelmingly preferred. She found that the Hong Kong accent was the object not just of low ratings but of ridicule: subjects laughed at the accent, even though they shared it. Luk explained this unusual finding by arguing that an institutionalised and socially accepted variety has not developed in Hong Kong because the ethnically homogeneous population have no need for English as a lingua franca or for in-group solidarity. Luk and Tauroza's subjects (schoolchildren) are much more representative of the typical product of the Hong Kong education system than are the speakers used by Bolton and Lim (2002).

Li (1999) argues similarly that 'there is no societal basis for a nativized variety of ''Hong Kong English'' ' (1999: 95), echoing Luke and Richards' (1982) view that the norms of correctness for the principle uses of English in Hong Kong, education, law, government, and business are all exonormative. These papers lend no support to the claims made in Bolton and Kwok's (1990) study and show attitudes directly opposed to those of the Indian and Singaporean subjects reported by Kachru (1992).

Conclusion

I have identified pronunciation errors which, at least from the perceptual standpoint of a native listener, involve phonemic overlap. These errors have been documented in previous studies but have not previously been identified as co-occurring within single words as they are repeatedly found to do in this study. These co-occurrences have the effect of making one word sound like another word, or perhaps like a nonsense word; the meaning of the resulting utterance, if there is any, is a matter of chance.

Although overt signalling of breakdowns in communication scarcely occurred in the data, it is argued that lack of overt comment by interlocutors on another's inadequate language use is not evidence of satisfactory communication; there are powerful social factors which mitigate against such overt comment.

It is argued on phonological grounds that instability of the accent, the repeated co-occurrences of phonemic overlap in the data, and the fact that for the most part the pronunciation is clearly due to transfer from Cantonese, all undermine the attempt to establish a 'phonology of Hong Kong English', whether this be intended merely as an academic exercise or whether it is intended to be applied in the classroom, as Mohanan clearly intended for the Singaporean context.

Turning to sociolinguistic issues, it is argued that related attempts to establish a 'sociolinguistic space' for Hong Kong English as a legitimate variety analogous to Singapore English have been based on an intellectual elite which is not representative of grass-roots Hong Kong speakers. These attempts have ignored both attitudinal studies showing the low esteem in which Hong Kong people themselves hold the accent and the principal uses to which English is put in Hong Kong, all of which point to exonormative standards of correctness. The analogy with Singapore is a false one, for the two places are sociolinguistically entirely different.

In certain informal situations, none of the errors documented here might matter. Indeed, it seems that the learners managed to perform the 'communicative activities' assigned to them despite grossly inaccurate pronunciation of key words. But this is less evidence of acceptable pronunciation than of the fact that such activities may be poor practice for language use outside the classroom.

In high stakes situations, on the other hand, errors such as these might well contribute to an unfavourable impression of the speaker and thus to professional or personal disadvantage. In conversations with non-native speakers from another language background, the effect of these coupled with the listeners' first language transfer effects could result in even more difficulty.

The maintenance of such extreme localised idiosyncracies is an unwarranted imposition on the patience of listeners in an international context.

I believe it is clear that the errors found in the present data are sufficiently severe to justify urgent attention to raising pronunciation standards and that this is not a matter either of slavishly aping former colonial masters, nor of unreasonable insistence on perfect RP, nor of exterminating local character. It is a matter of hard-pressed learners, whose first language causes severe difficulties in pronouncing English, trying to maintain intelligibility, often in the face of inadequate instruction and feedback.

Hongkongers have the right to an education which enables them to make use of English as they wish. None of the major uses of English in Hong Kong require the development of an in-group localised dialect: they all point the way to English for international communication. To help these subjects to take their role confidently on the world stage, an urgent attention to clear modelling and teaching of pronunciation is needed.

Correspondence

Any correspondence should be directed to Richard Stibbard, School of Arts, University of Surrey, Guildford, GU2, 7XH, UK (rmstibbard@yahoo.co.uk).

References

Boersma, P. and Weenink, D. (2003) *Praat: Doing phonetics by Computer*. Version 4.0.11. Computer program.

Bolton, K. (2000) The sociolinguistics of Hong Kong and the space for Hong Kong English. *World Englishes* 19 (3), 265–285.

Bolton, K. (2002) Chinese Englishes: From Canton jargon to global English. *World Englishes* 21 (2), 181–199.

Bolton, K. and Kwok, H. (1990) The dynamics of the Hong Kong accent: Social identity and sociolinguistic description. *Journal of Asian Pacific Communication* 1 (1), 147–172.

Bolton, K. and Lim, S. (2002) Futures for Hong Kong English. In K. Bolton (ed.) *Hong Kong English* (pp. 295–313). Hong Kong: Hong Kong University Press.

Boomer, D.S. and Laver, J.D.M. (1968) Slips of the tongue. *British Journal of Disorders of Communication* 3, 2–12.

Brazil, D. (1997) *The Communicative Value of Intonation in English* (new edn) Cambridge: Cambridge University Press.

Brown, A. (1995) Minimal pairs: Minimal importance? *English Language Teaching Journal* 49 (2), 169–175.

Brown, A. (2000) Tongue slips and Singapore English pronunciation. *English Today* 16 (3), 31–36.

Chan, A.Y.W. and Li, D.C.S. (2000) English and Cantonese phonology in contrast: Explaining Cantonese ESL learners' English pronunciation problems. *Language, Culture, and Curriculum* 13 (1), 67–85.

Hung, T.T.N. (2000) Towards a phonology of Hong Kong English. *World Englishes* 19 (3), 337–356.

Hung, T.T.N. (2002) Towards a phonology of Hong Kong English. In K. Bolton (ed.) *Hong Kong English* (pp. 119–140). Hong Kong: Hong Kong University Press.

Jenkins, J. (2000) *The Phonology of English as an International Language – New Models, New Norms, New Goals*. Oxford: Oxford University Press.

Jenkins, J. (2002) A sociolinguistically based, empirically researched pronunciation syllabus for English as an International Language. *Applied Linguistics* 23 (1), 83–103.

Kachru, B.B. (1992) *The Other Tongue. English Across Cultures* (2nd edn). Urbana, IL: University of Illinois Press.

Li, D.S. (1999) The functions and status of English in Hong Kong: A post–1997 update. *English World-wide* 20 (1), 67–110.

Luk, J. (1998) Hong Kong subjects' awareness of and reactions to accent differences. *Multilingua* 17 (1), 93–106.

Luke, K.K. and Richards, J.C. (1982) English in Hong Kong: Functions and status. *English World-wide* 3 (1), 47–64.

Mohanan, K.P. (1992) Describing the phonology of non-native varieties of a language. *World Englishes* 11 (2/3), 111–128.

Peng, L. and Setter, J. (2000) The emergence of systematicity in the English pronunciations of two Cantonese-speaking adults in Hong Kong. *English World-wide* 21 (1), 81–108.

Tauroza, S. and Luk, J. (1997) Accent and second language listening comprehension. *Regional English Language Centre Journal* 28, 55–71.