March, 2009

World Englishes, English as a Lingua Franca and the case of Hong Kong English

Andrew SEWELL, Lingnan University, Hong Kong

Available at: https://works.bepress.com/asewell/2/
World Englishes, English as a Lingua Franca, and the case of Hong Kong English

ANDREW SEWELL

Introduction

Mario Saraceni’s article (English Today 94) provides some useful insights into the current debates about English as a Lingua Franca (ELF). His discussion of the background to this debate identifies three viewpoints: a traditional ENL view with its adherence to native-speaker models; the WE (World Englishes) paradigm with its ‘pluralised and pluricentric view of English in the world’; and the emerging ELF position, with its rejection of native-speaker norms in favour of ‘endonormative realisations of lingua franca varieties’ (Alessa Cogo, English Today 95). However, Cogo believes that the second and third positions are not separate paradigms, and that ELF sits ‘comfortably within a WE framework’, as claimed by Jenkins (2007:17). In this article, I would like to show how the two positions can work together to inform pedagogy by exploring the possible options for English pronunciation models in Hong Kong.

The Hong Kong background

Hong Kong is well known for its strongly exonormative orientation towards native-speaker models. This is an environment in which the ‘traditionalist’ position is dominant in English-language teaching. In the area of phonology, there have been descriptions of Hong Kong English (HKE) from a WE perspective (for example, Hung, 2000; and Deterding et al., 2008). But public acceptance or awareness of a local model is limited, and most discussion of HKE has been limited to academic discourse. Luk (1998:103) believes that ‘the easy availability of an idealised exonormative model of English, coupled with the high status attached to it has made the development of an institutionalised endonormative model unnecessary’. However, the shortcomings of the traditionalist, NS-centred position are clearly visible in Hong Kong. Some local examinations display a strongly exonormative view of accents, for example the Language Proficiency Aptitude Test (LPAT) which is aimed at ‘benchmarking’ the language proficiency of local English teachers. In an article about the development of the LPAT, Coniam and Falvey (2002:23) quote the following as an example of a descriptor of pronunciation at the ‘well above the benchmark’ level: ‘Pronunciation is completely error-free with no noticeable L1 characteristics...[a]ny
mistakes that occur can be categorised as lapses rather than systematic errors.’

This exemplifies the prevailing traditional view, namely that ‘error-free pronunciation is also accent free’ (Luk and Lin, 2006:10; italics in original). However, while this position is clearly untenable, there is a lack of agreement regarding what an acceptable and intelligible local model might be like. This leads Kirkpatrick (2007a:387) to call for a ‘codified description of the local bilingual variety of English at the phonological, lexical-grammatical and discourse-pragmatic levels, as exemplified by highly proficient users of English who are mother-tongue speakers of Cantonese’. The importance of using a high-proficiency model is also noted by the local scholars Luk and Lin (2006:17), who remark on the need to distinguish between ‘local accents and careless speech’.

The important issue, therefore, is to decide which local accent features are acceptable in a local model of English that can inform language use, especially in the areas of teaching and testing. One way of approaching the issue is to combine the insights of the WE and ELF paradigms, in order to deal with the shortcomings of the prevailing traditionalist view.

Descriptions of Hong Kong English phonology

The contribution of the WE position has been to provide descriptions of HKE phonology, which as far as possible ‘needs to be investigated on its own terms’ (Hung 2000:138). Recent studies include those by Hung (2000) and Deterding et al. (2008). Hung’s study involved recordings of 15 university students reading word lists, while that of Deterding et al. used interviews to explore the ways words are pronounced in connected speech. The subjects in the latter study were teacher trainees and their overall proficiency level appears to be higher than the first-year arts and science undergraduates in Hung’s study.

The main features of the Hong Kong accent found in both studies included a reduced vowel system, with a particular tendency to merge the vowel pairs /æ, e/ and /ɪ, iː/; such reduced systems are a common feature of so-called New Varieties of English or NVEs, for example Singapore English. Diphthongs such as /ei/ and /ær/, on the other hand, tended not to be monophthongized as in other NVEs such as Singaporean or Indian English which have a simpler inventory of true diphthongs (Hung, 2000:127). In terms of consonant sounds, the voiceless dental fricative /θ/ was used by a majority of the subjects in both studies. Deterding et al.’s data show that there was an exception to this in final position, where a majority substituted /I/. The voiced dental fricative /ð/ was assumed to be non-existent in HKE by Hung, with words such as this being pronounced [dis]; Deterding et al.’s study does not cover them in detail, but does mention the common substitution of [d]. Dark [I] vocalisation was found to occur in both studies. The HKE phenomenon of [n, l] conflation occurred quite widely in Hung’s data, but was found to be ‘rare’ by Deterding et al. Finally, Deterding et al. also looked at consonant cluster patterns and found that initial clusters were subject to substitution, with deletion often observed in final clusters. A summary of the similarities and differences between the two studies is shown in Table 1.

While the similarities between the two descriptions allow some generalisations to be made about what constitutes segmental phonological features of an ‘identifiable Hong Kong accent’ (Hung, 2000:119), the differences point to what might be called the variability problem: individual speakers vary greatly in the extent to which they display these features in their normal speech, so it is perhaps difficult to talk of a ‘Hong Kong accent’ without being clear about the proficiency level of the speakers. It seems likely that the subjects of the Deterding et al. study generally had a higher level of proficiency, as their phonological systems were somewhat more developed (there was less conflation of [n] and [l], for example). And while there will be features in common between speakers at various levels, there are very great differences between the Hong Kong accent of a high-proficiency speaker such as Chief Executive Donald Tsang and that of a low-proficiency speaker, such as the sole Hong Kong sample in the online Speech Accent Archive (Weinberger, 2008).

Researchers in the WE paradigm have acknowledged this variability, so that Hung (2000:122) visualises ‘a continuum with an “idealized” HKE phonology at one end and a standard British or American phonology...at the other’. But the variability problem is not insurmountable, and the insights provided by ELF research may be useful in evaluating
language data from a WE perspective. By doing so, the data can be ‘considered in pragmatic rather than a priori terms’ (Nelson, 1995:273).

**ELF insights into intelligibility**

As noted by Saraceni (2008:21), the work of Jenkins (2000) has been seminal in ELF research. The *Phonology of English as an International Language* not only identified the phonological features of English that were most likely to contribute to intelligibility problems in international communication, but also demonstrated ‘how things could be done differently’ (Saraceni, ibid.). For example, Jenkins’s list of core features provides an indication of how the features of native-speaker accents may also hinder intelligibility. In a much-quoted passage she talks of how native speakers will have to learn how to make themselves more easily understood in international contexts, so that ‘for those who have already reached adulthood it will be necessary to attend adult EIL classes, in the same way that “NNS adults” do’ (2000:228). Jenkins prefers the term ‘ELF’ to ‘EIL’, English as an International Language, in her later work, for example Jenkins, 2007.

The ‘core’ areas identified by Jenkins (2000:159) on the basis of empirical research into the factors affecting international intelligibility are listed below:

1. **Consonant sounds**: the inventory of permissible sounds resembles existing NS models, with the following differences:
   - Rhotic ‘r’ is preferred, so that RP speakers will need to modify this aspect of their speech.
   - The intervocalic [t] of RP is preferred to the voiced alveolar flap [ɾ] of GA in words like *better*.
   - Most substitutions of the dental fricatives (for example, [t] for [θ] and [d] for [ð])

---

**Table 1: HKE phonological features**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Vowels</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Merger of vowel contrasts,</td>
<td>Yes (seven monophthongs; [æ, e] and</td>
<td>Yes: merger of [æ, e] and [i, i:]</td>
</tr>
<tr>
<td>reduced vowel system</td>
<td>[i, i] merged)</td>
<td></td>
</tr>
<tr>
<td>FACE and GOAT ((ɛr) and (aɪ))</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>as diphthongs</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Consonants</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voiceless TH in initial position</td>
<td>[θ] used by ‘at least half’</td>
<td>[θ] for 64% of tokens (33% used [I])</td>
</tr>
<tr>
<td>Voiceless TH in medial position</td>
<td>[θ] used by ‘at least half’</td>
<td>[θ] for 60% of tokens (20% used [l])</td>
</tr>
<tr>
<td>Voiceless TH in final position</td>
<td>[θ] used by ‘at least half’</td>
<td>[θ] for 22% of tokens (67% used [f])</td>
</tr>
<tr>
<td>Voiced TH</td>
<td>Does not exist in HKE; substituted by [d]</td>
<td>Not reported in detail</td>
</tr>
<tr>
<td>Conflation of [n] and [l]</td>
<td>Yes (up to 37% of tokens)</td>
<td>Rare</td>
</tr>
<tr>
<td>[v] substitution [w]</td>
<td>Yes – [v] does not exist in HKE and is substituted by [w] or [f]</td>
<td>Not reported</td>
</tr>
<tr>
<td>Dark [l] -vocalisation</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Consonant clusters</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial</td>
<td>Not studied</td>
<td>[kr] clusters often realised as [kl], e.g. crowded [klaudl]; [pr] clusters less so</td>
</tr>
<tr>
<td>Final</td>
<td>Not studied</td>
<td>Deletion in 54% of cases before a pause or before a word beginning with a vowel</td>
</tr>
</tbody>
</table>

---

WORLD ENGLISHES, ENGLISH AS A LINGUA FRANCA, HONG KONG ENGLISH 39
are permissible. Pronouncing three as [triː] and that as [dæt] are seen as unproblematic. The same applies to dark [l] substitutions such as [wiʊ] instead of [wi] for will.

- Close approximations to core consonant sounds are permissible while certain L1-influenced substitutions are not, for example the pronunciation of very as [βeri] by Spanish learners, which tends to be heard as berry.

2 Phonetic requirements:
- The plosives [p, t, k] must be aspirated in initial position, to avoid confusion with 
  [b, d, g].
- The phenomenon of ‘pre-fortis shortening’ whereby the vowel in cap is shorter than the vowel in cab, whilst being phonemically equivalent, should be retained.

3. Consonant clusters:
- Initial clusters should not be simplified, so that the Hong Kong English pronunciation of produce as [pudʒuːs] (Chan and Li, 2000:82) would not be permitted.
- Medial and final clusters can be simplified, but only according to L1 rules of elision. This means that [t] and [d] may be deleted from words with three-consonant clusters like facts and pounds, or from two-consonant clusters where the next word begins with a consonant (such as strict rules).

4 Vowel sounds:
- Vowel length (quantity) contrasts should be maintained.
- L2 vowel qualities are permissible if consistent, but [ə] should be preserved.

5 Nuclear (tonic) stress: nuclear stress placement and the division of speech into ‘word groups’ (also called tone units, or intonation units) are important for intelligibility.

While there are specific examples of the possible modifications that may be required from native speakers, one of the implications of Jenkins’s Lingua Franca Core (LFC) is that an internationally intelligible English accent will not be easy to attain; while some features can be considered as candidates for removal from pronunciation teaching syllabi, others appear to add to the complexity of the task, rather than reducing it. So while dental fricative substitutions are seen as non-threatening for international intelligibility, the requirements for aspirated [p, t, k] would actually add to the task for those whose accents tend to feature reduced aspiration of these sounds (for example, speakers of Indian English or Yorkshire English; see Wells, 1982).

The intelligibility filter

It is possible to evaluate the specific features of the Hong Kong accent by passing them through the ‘intelligibility filter’ provided by ELF research: this procedure resembles the technique outlined in Brown (1991), but with reference to empirical data on intelligibility. Of the features identified by Hung and Deterding et al., the following would at first sight seem to be unproblematic for international intelligibility: the dental fricative substitutions (usually [f] for [θ] and [d] for [ð]), dark [l] vocalisations and the merging of the [æ] and [e] vowels. While I do not wish to examine in detail the arguments regarding the relative importance of these features, it is worth considering each of them briefly.

Dental fricative substitutions were also noted in the study of ASEAN English carried out by Deterding and Kirkpatrick (2006), and they too were found not to affect mutual intelligibility. They are also present in many L1 English accents, as noted by Jenkins (2000, 2007). Although it is debatable how far L1 accent features should be allowed to inform an assessment of international intelligibility – especially if it is accepted that the number of non-native speakers of English exceeds the number of native speakers – it does seem that there is something in language universals that militates against an insistence on L1 English norms. Features such as these are termed ‘weak links’ of the target language by Williams (1987), who believes they may be important factors in language acquisition and change. Psycholinguistic principles also suggest the relative unimportance of the dental fricatives, as the small number of minimal pairs involving these sounds results in a low probability of lexical confusion: ‘in terms of minimal pairs, all pairings involving [θ, ð] are unimportant’ (Brown, 1991:121).

Similar arguments can be raised in favour of ignoring dark [l] vocalisation, which also occurs frequently in L1 accents (as well as in
NVEs such as Singaporean English). Bauer (1986) found that in New Zealand English, vocalised [I] was so prevalent that many people could not make a dark [I] preconsonantally (in Shockey 2003:35). Interestingly, such substitutions may also be indications of language change on a wider scale. Shockey (2003:112) notes that modern Portuguese shows final [I] vocalisation to be 'a living process': Spanish *mal*, Portuguese *mau*.

Although merging the [ae] and [e] vowels would not appear to cause intelligibility problems according to Jenkins's Lingua Franca Core, this is a slightly more complex issue. Firstly, it is not entirely clear what Jenkins (2000:159) means by saying that 'L2 regional vowel qualities are permissible'. Of course, there is considerable variation in vowel quality across different speakers and different accents. But while individual vowel qualities may vary, what is important for international intelligibility is that most vowel contrasts are maintained. Thus while a speaker's actual vowel qualities will show some divergence from the idealized versions implied by a vowel quadrilateral, if they are merged then what Brown (1991:77) calls 'the maintenance of perceptual distance' will be lost. This will almost certainly affect intelligibility.

Of course, this is not to suggest that vowels must necessarily be realised in native-like ways; it is not inconceivable that future varieties of English might distinguish between word pairs such as *pen* and *pan* somewhat differently. But there will still be a contrast between them, barring major changes to the lexicon. In addition, there may be ways in which Hong Kong listeners are better able to distinguish sound contrasts made by Hong Kong speakers, so that theoretical merging may not always mean actual merging. This was suggested by the data in Hung (2000:126), where Hong Kong subjects were better able to differentiate words containing the [u, ur] contrast, when read by a Hong Kong speaker, than native-speaker listeners were. However, if international intelligibility is accepted as being important – and it certainly is in Hong Kong, where English is principally a tool for international communication – then accepting the merger of [ae] and [e] would seem to be problematic. The merger of [I] and [i] caused intelligibility problems in Jenkins's data. Although this is not discussed by Jenkins herself, the importance of this contrast may be related to the relatively large number of minimal pairs involving this contrast (471, according to Higgins, 2008). The importance of such 'functional load' calculations is controversial, but the merger of [ae] and [e] – 302 pairs (Higgins, ibid.) – might also be expected to be problematic. Brown (1991) assigns top priority to this contrast, partly because of his experience of miscommunication involving near-minimal pairs such as *bag* and *peg* in Singapore.

I myself have experienced communication breakdown as a result of the [ae, e] merger in Hong Kong, as when a local artist told me that he was going to *send the painting*. Some ELF researchers might claim this native-speaker observation to be irrelevant, but if it causes problems for me it is likely to cause problems for others, assuming that auditory discrimination ability is by and large species-specific. It is not therefore entirely clear why Jenkins (2007:27) appears to propose 'the removal of perceptual distance' will be lost. This will almost certainly affect intelligibility.

Turning to the features of Hong Kong English that do not pass the test of international intelligibility, there is the conflation of [n] and [I] ('rare' in Deterding et al., but present in 37% of tokens in Hung) and the tendency to substitute [w] for [v] (also more prevalent in Hung's data). The simplification of initial clusters is proscribed by Jenkins's core, so that pronouncing *crowded* as *[klaʊdd]* would be expected to lead to intelligibility problems. The speakers in Deterding et al. deleted final consonants in 54% of clusters occurring before a pause or before a word beginning with a vowel, environments where deletion would not normally be expected in NS varieties of English (Schreier, 2005:30, in Deterding et al., 2008:157). This is therefore a feature that might lead to intelligibility problems according to Jenkins's criteria, in which final clusters can only be simplified according to L1 rules of elision.

But on further examination, some of the simplifications noted by Deterding et al. seem unlikely to be problematic, such as the common Hong Kong pronunciation of *I think* without the final [k] consonant. The relatively lower importance of final clusters is entirely
predictable, from a psycholinguistic standpoint: ‘The deletion of cluster-final segments at the end of lexical items results in a lower degree of information loss than reduction in the beginning of words’ (Schreier, 2003:12). There is a need for more research into the consequences for intelligibility of certain kinds of non-native cluster simplification. For example, in Hong Kong one can often hear multisyllabic words such as government or consultant produced without the final consonant; while native speakers would not normally do this it seems unlikely to affect intelligibility.

Other features of Hong Kong English may be found to actually contribute to intelligibility, rather than detract from it, especially in the suprasegmental realm. The tendency towards syllable-timed rhythm noted by Deterding et al. has been argued by Kirkpatrick (2007b) to enhance intelligibility, although this remains controversial and relatively unsupported by empirical evidence.

**Discussion**

The above combination of approaches and data has generated some interesting insights into both the WE and the ELF perspectives. One possible interpretation is that while the ENL model certainly has limitations, passing the features of a local variety through what I have called the ‘intelligibility filter’ results in a possible local pronunciation model that is not very significantly different from the existing native-speaker models. However, it is equally true to say that there are prevalent features of the Hong Kong accent that are unlikely to affect intelligibility and should therefore be given a lower priority in teaching, as well as being brought to the attention of those involved in language testing.

Naturally, in a full-scale evaluation of pronunciation models there would be far more to consider than intelligibility alone. It is worth questioning the extent to which considerations of international intelligibility should be employed at all in evaluating the features of local accents. After all, one of the objectives of ELF research has been to explore ways in which L2 learners of English can retain their L1 identity; too intense a focus on intelligibility may even have the effect of discouraging the use of local varieties. There remains a certain tension between the ‘common core’ of features suggested by ELF research and the ‘multiplicity and creativity’ of local varieties perceived by the WE movement, as noted by Berns (2008: 328). Another factor to take into account would be learnability, the relative difficulty for learners of acquiring certain features.

The separate dimension of acceptability would also be important, although far less amenable to empirical investigation. In considering dental fricative substitutions, for example, the conclusion of Deterding et al. (2008:153) is that these sounds are ‘something of a shibboleth in new varieties of English’. Perhaps because the voiced dental fricative [ð] occurs so frequently in English, it is likely that listeners’ perceptions of a speaker are strongly affected by their pronunciation of this sound. Rightly or wrongly, such substitutions are stigmatized by many. Applying the ‘precautionary principle’ to syllabus design and teaching leads to the conclusion that if these sounds are given a lower priority, learners should be made aware of the possible disadvantages of not acquiring them. One desirable outcome might be a ‘dual competence’ in which English users manipulate features according to whether they wish to express their local or global selves, giving learners what Shaw (2008:49) calls ‘the freedom to choose one’s belongingness’.

Finally, one advantage of combining the WE and ELF perspectives and adopting an intelligibility-oriented approach is that it may serve to counter the perception that local models hinder global intelligibility. While language attitudes may be deeply entrenched, and while they need to be taken seriously, it is important to remember that ‘acceptability’ is not an immutable concept; there may be a sense in which acceptability can be engineered (Bamgbose, 1998: 4). If local stakeholders – learners, parents, teachers, principals and bureaucrats – can be persuaded that a local accent exists, and that certain of its features do not threaten international intelligibility, then it will be easier to overcome the prevailing belief that the local variety is merely a collection of errors.

**References**


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