University of Reading

Word juncture characteristics in world Englishes: a research report

Article

Accepted Version

Setter, J. ORCID: https://orcid.org/0000-0001-7334-5702, Mok, P., Low, E. L., Zuo, D. and Ao, R. (2014) Word juncture characteristics in world Englishes: a research report. World Englishes, 33 (2). pp. 278-291. ISSN 1467-971X doi: https://doi.org/10.1111/weng.12085 Available at https://centaur.reading.ac.uk/54120/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.
Published version at: http://onlinelibrary.wiley.com/doi/10.1111/weng.12085/full
To link to this article DOI: http://dx.doi.org/10.1111/weng.12085

Publisher: Wiley

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

University of Reading

Reading's research outputs online

The Perception of Word Juncture Characteristics in Three Varieties of English: A Research Report

Authors

Prof Jane Setter (correspondence)
Department of English Language & Applied Linguistics
University of Reading
Whiteknights, PO Box 218
Reading, UK
RG6 6AA
Email: j.e.setter@reading.ac.uk

Dr Peggy Mok
Department of Linguistics and Modern Languages
Chinese University of Hong Kong
Hong Kong
Email: peggymok@cuhk.edu.hk

Dr Ee Ling Low
Department of English Language and Literature
National Institute of Education
Singapore
Email: eeling.low@nie.edu.sg

Ms Donghui Zuo
Department of Linguistics and Modern Languages
Chinese University of Hong Kong
Hong Kong
Email: donghuizuo@gmail.com

Mr Ran Ao
Office of Teacher Education
National Institute of Education
Singapore
Email: ran.ao@nie.edu.sg

The Perception of Word Juncture Characteristics in Three Varieties of English: A Research Report

Abstract

The subtle juncture cues in older varieties of English such as Received Pronunciation can be difficult for speakers of new English varieties to perceive. This study looks at the perception of word juncture characteristics in three varieties of English (British, Hong Kong and Singapore) amongst British, Hong Kong and Singaporean listeners in order to widen our understanding of English juncture characteristics in general. We find that, even though reaction time data indicates that listeners perform quickest in the variety they are most familiar with, not only are juncture differences in British English difficult for Hong Kong and Singaporean listeners to perceive, they are also the most difficult for British listeners. Juncture characteristics in Hong Kong English are the easiest to distinguish among the three varieties.

Introduction

This study investigates the perception of word juncture characteristics in three varieties of English: Hong Kong English (HKE), Singapore English (SE) and British English (BE). Juncture refers to 'any phonetic feature whose presence signals the existence of a grammatical boundary' (Trask 1996:189). In more general terms, it means the boundary between two syllables. In this paper we are looking at the boundary in ambiguous word pairs, e.g., *great eyes* versus *gray ties*, in which both phrases have the same phonemic representation – /greɪtaɪz/ – but different patterns in allophonic variation across the word boundaries.

Understanding connected speech demands that the listener identifies where words begin and end, but in fluent speech there is no obligatory gap between words to signal this information. However, people are usually able to understand speech and discern individual words using a combination of contextual information and subtle cues in the speech signal. Studies of word juncture characteristics in older varieties of English (OVEs, e.g., British or American English) abound, but virtually no work has been done on new varieties of English (NVEs) in East Asia. This study compares the perception of word juncture boundaries in two new East Asian English varieties (HKE and SE) with a traditional one (BE) to investigate the perceptual differences in these accents.

Juncture Characteristics

The seminal work on the production and perception of acoustic juncture cues in English was conducted by Lehiste (1960). She studied twenty-five pairs of words or phrases that are phonemically the same but have different juncture – e.g. *nitrate* versus *night-rate* (both phonemically /naɪtreɪt/) – and found that listeners could identify them correctly because there are different juncture cues to signal where the syllable boundaries fall. For

instance, the allophone of /r/ in *night-rate* is voiced but it is almost completely devoiced in *nitrate*. She concluded that there are regular juncture characteristics correlating with word/syllable boundaries. Since then, many studies of the production and perception of juncture cues have been conducted in English (e.g. Christie 1974; Nakatani and Dukes 1977; Mattys and Melhorn 2007) and in other languages – for example, Swedish (Gårding 1967), French (Rietveld 1980) and Dutch (Quené 1993). An increasing body of research shows that fine phonetic detail like that observed in juncture cues is systematic and provides useful linguistic information for the listeners (e.g. Hawkins 2003).

Locating syllable boundaries in English is not always straightforward, and the syllabification of intervocalic consonants in English has been a controversial issue. Many psycholinguistic and phonetic studies showed that syllabification of intervocalic consonants depends on several factors or principles; the maximal onsets principle, the sonority contour of a syllable, stress placement, vowel length, the phonotactic legality of the sequences, the phonetic identity of the consonants and morphological structure of the words can all play a part (see, e.g., Boucher 1988; Redford and Randall 2005; Treiman and Danis 1988; Treiman and Zukowski 1990). These principles vary in importance and may result in different syllabification of the same sequence.

Many acoustic juncture cues contours have been identified by the studies cited above, such as the presence of a glottal stop or laryngealization before initial vowels, variation in segmental duration, final lengthening, the presence of short pauses, allophonic variations (e.g., the presence of a stop burst), differences in formant transitions, fundamental frequency and intensity. However, not all of these phonetic properties are employed equally by listeners in determining syllable boundaries. For instance, Christie (1974) demonstrated that the presence or absence of formant transitions in synthetic speech does not significantly affect listeners' judgments of syllable boundaries in English. Moreover, Schwab, Miller, Grosjean

and Mondini (2008) showed how a listener's ability to use various juncture cues to identify word boundaries can be affected by speaking rate, as perception of juncture boundaries in fast speech was less accurate than in slower speech. Ultimately, therefore, it is necessary to study both production and perception in order to identify the crucial juncture characteristics.

In communicative situations, however, it is not usually the case that a listener has only phonetic information to rely on where the correct parsing of ambiguous word pairs or sequences are concerned, as context has a crucial role to play and can bias or override any acoustic cues available to a listener in a target utterance. Mattys and Melhorn (2007) showed how the use of a carrier sentence can strongly influence perception of phonetically similar sequences; for example, in the case of the sequence /plʌmpaɪ/, listeners were more likely to choose *plum pie* over *plump eye* if the carrier sentence was *The baker looked at the drawing of a …* as opposed to *The surgeon looked at the drawing of a …*, even if the phrase when presented on its own had been parsed as *plump eye* by listeners. This effect diminished somewhat if the context was not so obviously biased semantically, e.g., in a carrier sentence such as *The girl looked at the drawing of a …*

In English language teaching materials, OVEs such as Received Pronunciation (RP) are used as models in a variety of settings, and the language in recordings which accompany these materials is often hyper-articulated to varying degrees. Indeed, in a 'normal' communicative context it is usual for a speaker to adapt his or her speaking style to accommodate to the listener. Lindblom (1990) developed the Hyper- and Hypo-articulation theory (H&H) of speech production in order to account for the use of this strategy in normal communicative contexts, in which a speaker will use more or less articulatory effort dependent on their evaluation of the needs of the interlocutor, i.e., how easy or difficult it is for the listener to understand the message. This difficulty might be owing to factors such as lack of contextual cues, background noise, or the language proficiency of the listener. H&H

theory therefore places the burden of producing clear enough acoustic cues on the side of the speaker; recordings made for ELT materials could be seen as an institutionalised variant of this accommodation strategy. In spite of this, it is highly likely that the subtle juncture cues in OVEs will be difficult for speakers of new varieties of English (NVEs) and/or learner varieties to perceive in real communicative contexts and, therefore, a contributing factor in the ability to understand these speakers. However, it is possible that the reverse is also true – that speakers of NVEs will have speech patterns which differ enough from OVEs to make their speech a challenge to understand. This is taken up by Kim, Stephens and Pitt (2012: 509) when they comment that '[s]uccessful communication requires the listener to have learned how to segment speech from a variety of talkers (e.g., native, foreign-accented) speaking in a variety of styles (e.g., careful vs. casual speech)'. A literature search suggests that there is virtually no study investigating the perception of juncture in NVEs, although studies on the phonology of NVEs have been increasing in recent years; see, for example, Deterding, Brown and Low (2005) on SE, Hung (2000), Setter (2008) and Deterding, Wong and Kirkpatrick (2008) on HKE, Gargesh (2008) on Indian English, Tayao (2008) on Philippine English. This paper aims to move towards filling this gap by providing data on speakers of both NVEs (HKE and SE) and OVE (BE).

Hong Kong and Singapore Englishes

English is integrated into Hong Kong life in a way unusual to many other settings in which it is used (Setter, Wong and Chan 2010), and this has resulted in a thriving and developing variety of English. Although there are no, or very few, native speakers of HKE in the same way as there are native speakers of SE or Indian English, HKE is an identifiable English variety, which shows no signs of falling out of use, or becoming merely another learner variety.

In terms of the development of NVEs, Schneider (2007) put HKE at Phase 3 in his Dynamic Model of the Development of Postcolonial Englishes (2007), i.e., at 'nativization'. This indicates that the variety is in a state of 'cultural and linguistic transformation' (Schneider 2007: 40) and that there is a move towards independence from the distant country of origin politically, linguistically and culturally. In the context of the return of HK to China in 1997, Schneider noted that the drive to use English is 'stronger than might have been anticipated' (2007: 139), and, indeed, Li's (1999) description of English as a 'value-added language' in HK indicates the importance attached by Hong Kong people to having an excellent command of English.

As mentioned earlier, HKE has identifiable phonetic as well as other linguistic features which have been documented in a number of recent studies. Bolton and Kwok (1990), for example, provided a brief overview of some phonological features, and Chan and Li (1999) compared HKE with BE from a learner English perspective. The first paper to attempt a detailed and thorough account of the segmental phonetics and phonology of HKE in terms of its being an emergent NVE is Hung (2000) which, using a quantitative, acoustic methodology based on word list data, gave an inventory of phonemes, looked at the phonetic realization of those phonemes, and suggested distribution of phonemes in the syllable. This was followed up by Deterding et al.'s (2008) study which uses connected speech. Setter (2008) discussed the HKE syllable in depth, and Wong and Setter (2002) considered a possible /n/ and /l/ merger in syllable initial position. Studies on suprasegmental features include Hung (2005), Luke (2008) and Wong (1991, 2004) on word stress and Setter (2005) on rhythm.

In contrast with HKE, which is still moving towards NVE status, SE is a recognized NVE. Schneider (2007: 153) placed SE at Phase 4 in the Dynamic Model, 'endonormative stabilization', which means that there is political independence and cultural self-reliance in

the environment where the variety is used, and also indicates that the English spoken in the territory is a variety in its own right, on equal terms with other recognized varieties. There are many studies on SE phonetics and phonology, both acoustic and auditory (see the bibliographies compiled in Brown 2005 and Low and Azirah 2012). Using the National Institute of Education corpus of spoken Singapore English (NIECSSE) (Deterding and Low 2001) as its data source, Deterding et al. (2005) contained contributions on consonants, vowels, suprasegmentals, phonetic aspects of conversation analysis and intelligibility, and it is far from being the only publication on SE phonetics and phonology. One can claim that the phonology of SE has been thoroughly researched and defined.

Both HKE and SE are heavily influenced by Chinese languages. Comparing the phonology of two varieties with a similar linguistic background but different social status can provide new insights for both varieties. Deterding et al. (2008) found that, although Hong Kong English shares many phonological features with Singapore English, it also contains features found only in British English but not in other Englishes in South-East Asia – for example, the use of [f] for initial 'th' in content words and the fronting of [u]. Such differences were attributed to the different developmental stages of the two English varieties. Therefore, it is quite possible that differences in juncture characteristics can also be found in the two varieties, but so far no study has investigated them yet.

An interesting development in the description of HKE and SE is presented in Kirkpatrick, Deterding and Wong (2008), who found that educated HKE was highly intelligible to Singaporean and Australian listeners in comparison with SE, which had been tested in an earlier study (Kirkpatrick and Saunders 2005). In both studies, Kirkpatrick and colleagues played short excerpts of speech from several speakers to listeners and used worksheets containing comprehension questions to deduce the intelligibility of the speakers. Although the listeners did well in both varieties, it is interesting to note that HKE was found

to be more intelligible. It is possible that juncture cues played a part, but this was not investigated in any depth in either study.

The Present Study

This paper reports perception data on word-pair juncture characteristics in HKE, SE and (Southern Standard) BE collected from Hong Kong, Singaporean and British listener groups. Following review of the literature, our hypotheses are as follows:

1. Juncture boundaries in BE will be most difficult to distinguish for listeners in all three varieties;
2. Juncture boundaries in HKE will be the easiest to distinguish;
3. Hypothesis 1 notwithstanding, listeners will do best in terms of percentage correct and have the fastest reaction times in their own variety in comparison with the other groups.

Hypothesis 1 is proposed because of the subtle cues in allophonic variation and greater linking between word boundaries in BE reported in previous literature, which we feel will make it difficult for the non-BE listeners to perceive the difference between BE word pairs. The anticipated differences in these features in HKE and SE by comparison (e.g., HKE and SE both have strong glottal reinforcement of final stop consonants) should make certain HKE and SE pairs easier to distinguish. Despite the work of Kirkpatrick et al. (2008), it is unclear whether HKE and SE differ in clarity of juncture boundaries but, based on that study, Hypothesis 2 is put forward, i.e., HKE will be the easiest variety in which to discriminate between the word pairs. Finally, owing to a listener's experience with his or her own variety, it is hypothesized that the best performance in BE will be from British listeners (BLs, where abbreviated), the best performance in SE will be from Singaporean listeners (SLs), and the

best performance in HKE will be from Hong Kong listeners (HKLs), both in terms of percentage correct and reaction time (Hypothesis 3).

We confirm that the research meets the ethical guidelines of the institutions in which it was carried out.

Method

Participants. 60 listeners took part in the study. 25 listeners for each of HKE and BE were recruited in Hong Kong and Reading, UK respectively. It was initially intended to recruit 25 listeners in Singapore, but ultimately only 10 participants took part.

Participants were all university students with no speech or language impairment. The British listeners were all monolingual native English speakers, although some of them had had limited exposure to foreign languages. The Hong Kong and Singapore listeners had not lived in another English-speaking country, and had received all of their education in these two places respectively. The Hong Kong and Singapore participants were paid to participate in the perception experiment, whereas the British participants participated in the experiment for course credit.

Materials. 24 juncture boundary pairs adapted from lists used in previous research (Lehiste 1960; Schwab et al. 2008) were used in this study (see Table 1, Appendix). 20 pairs involve a single consonant at the juncture boundaries, either a stop (e.g., *wipe ink* versus *why pink*) or a sonorant (e.g., *no notion* versus *known ocean*) while four pairs involve a consonant cluster (e.g., *my train* versus *might rain*).

As this was a pilot study, in order to keep the perception experiment to a manageable size, one typical female speaker from each variety was recorded reading the materials several times. The three speakers were born, had grown up and been educated to tertiary level in the

UK, Hong Kong and Singapore respectively. The HKE and SE speakers had not spent any time in other English-speaking countries. The BE speaker was a 43-year old academic whose voice has been used in commercial recordings for standard British English materials requiring a near-RP accent. The HKE speaker was 25 years old and the SE speaker was 40 years old at the time of recording. They were confirmed as representative of speakers of educated HKE or SE by other speakers of that accent.

The 24 juncture pairs were put in a carrier phrase 'HE writes _____.' with emphatic stress falling on the word 'HE' in order to minimize stress difference between the two target words in the juncture pairs. Each one was read three times. Subsequent careful auditory and visual inspection of the sound files using PRAAT were carried out to select the tokens with comparable degree of stress. The pairs were then excised from the carrier phrase for the perception experiment, i.e., only the word pairs themselves were presented to participants. The target words were cut from the burst release (if the initial segment was a stop), the beginning of friction (if the initial segment was a fricative or affricate), or the beginning of voicing (if the initial segment was a sonorant).

Procedure. The perception experiment is an identification task. The software DMDX (available from http://www.u.arizona.edu/~kforster/dmdx/dmdx.htm) was used, running on a laptop or desktop computer. The software allowed us to collect both accuracy (%C) and reaction time (RT) data. The participants heard a recording (e.g., *wipe ink*) and saw two sequences on the screen (e.g., *wipe ink* and *why pink*) presented simultaneously and several centimetres apart, one on the right and one on the left. They then made their choice by pressing the 'z' key for the sequence on the left of the screen or the 'm' key for the sequence on the right. The positions of the target sequences were counterbalanced, i.e., the match between the audio stimuli and the position of the correct item on screen for selection were equalized. RT was tracked from the beginning of the sound file. The time-out time was

8000ms, i.e., if the participant did not respond within 8000ms the programme moved on to the next pair automatically.

Before the actual experiment, a practice session with 10 juncture pairs which did not appear in the main study was given to familiarize the participants with the experimental procedure.

The experiment was divided into three sections – one section for each variety – and each section was divided into four blocks. Rests were given between sections and between blocks. 96 tokens (24 pairs × two target sequences × two positions) were used for each variety, with 288 tokens (96 tokens × three varieties) in total for the perception experiment. The tokens were randomized within blocks for each participant. Stimuli blocks were presented in order of anticipated difficulty. The HKE participants heard the blocks in the order HKE – SE – BE and the SE participants heard them in the order SE – HKE – BE. We allowed them to hear the accent they were most familiar with first as we anticipated they would do better in this accent, and we scheduled BE last as we anticipated it would be the most difficult. For the BE participants, half of them heard the stimuli blocks in the order HKE – SE – BE, and half heard them in the order SE – HKE – BE. We again presented the BE last but, as we were not able to anticipate how they would react to the other two varieties, we decided to counter-balance the presentation of them.

Statistical significance level is set at $p \leq 0.05$. As there is a choice of two sequences for each auditory stimulus, chance level is 50%.

Results

Table 2 in the Appendix gives an overview of accuracy in terms of the average percentage correct, maximum and minimum score and the standard deviation for each of the listener groups respectively, and Table 3 presents the RT results, also in terms of average RT, maximum and minimum RT and standard deviation. Statistical significance is indicated on

these tables. Figures 1, 2 and 3 show the average percentages of correct identification and RTs in the three varieties for Hong Kong, British and Singapore listeners respectively.

Concerning percentage correct (%C), each set of listeners scored best on HKE, followed by SE, with BE as the lowest scoring variety. One-way ANOVAs and post hoc Bonferroni tests confirm that the differences between varieties are all significant for Hong Kong listeners ($p<0.001$) (Figure 1, top). For British listeners (Figure 1, centre), the difference is significant between BE and HKE ($p<0.001$), and between SE and HKE ($p<0.01$), but there is no significant difference between BE and SE ($p=0.701$). Where SE listeners were concerned (Figure 1, bottom), their patterns are similar to HKE listeners in that the differences are significant among all varieties: BE vs. SE ($p<0.001$), BE vs. HKE ($p<0.001$) and SE vs. HKE ($p<0.01$). Although the RT data show that the listeners responded the quickest in their own variety, there is no significant difference in RT among varieties for any group of listeners, although in the case of BE listeners it approaches significance between BE and SE ($p=0.058$). The maximum RT was 2411ms, i.e., no participant failed to answer within 8000ms.

On the BE task, the British listeners did best with an average percentage score of 71.46%, outperforming both the Hong Kong (59.58%, $p<0.001$) and the Singaporean listeners (64.48%, $p<0.05$), and the Singaporean listeners outperformed the Hong Kong listeners ($p<0.01$) on the same task. The Hong Kong listeners outperformed the British listeners on both the HKE and SE tasks with a significance level of $p<0.001$; Hong Kong listeners scored 89.5% on average on HKE and 78.33% on SE in comparison with the British listeners' 81.7% and 74.17% respectively. The Singaporean listeners did best in HKE, with an average 86.35%, but this unexpected finding was not statistically significant in comparison with the performance of the other listeners. No participant scored lower than 51% in any one task, i.e., they all performed at above chance level (strictly speaking).

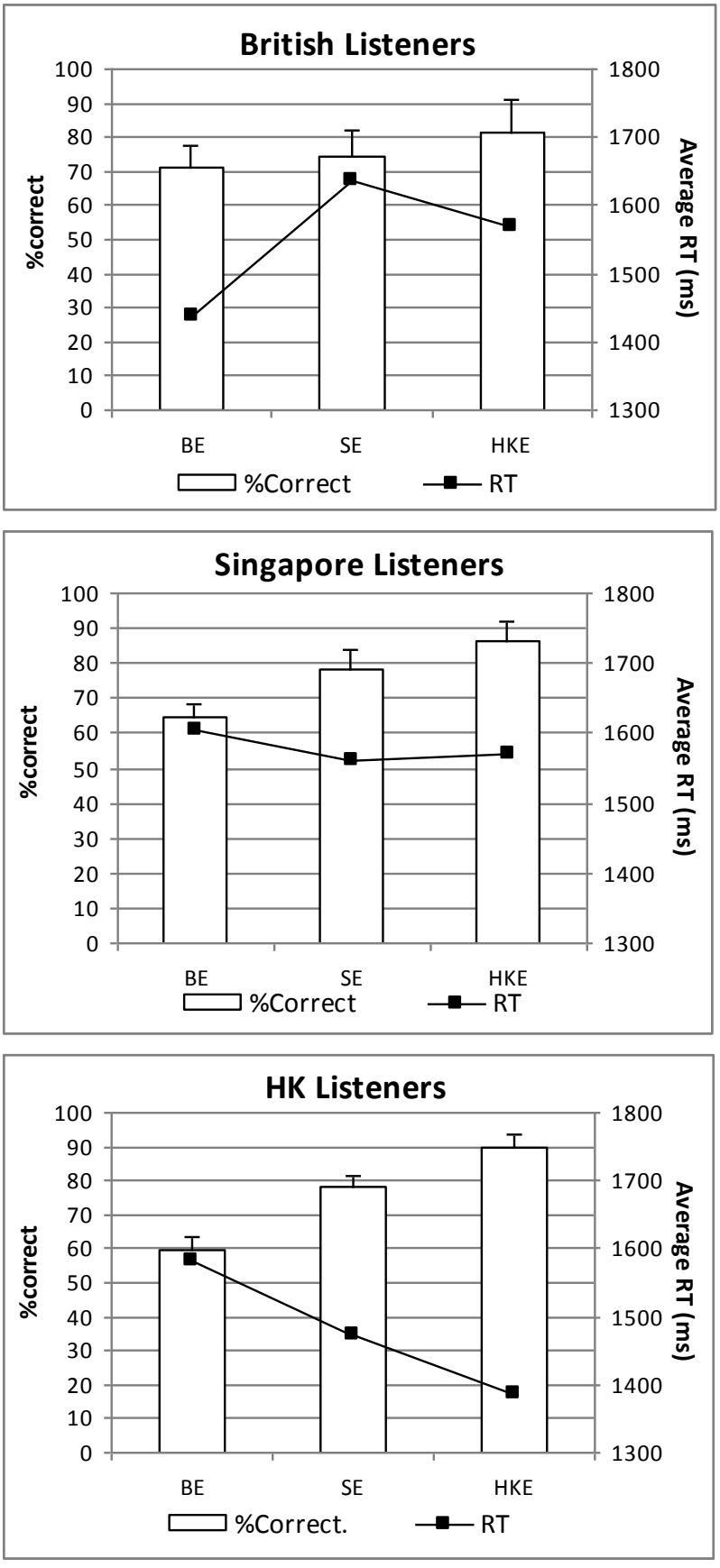


Figure 1: %C and RT for the three English varieties.

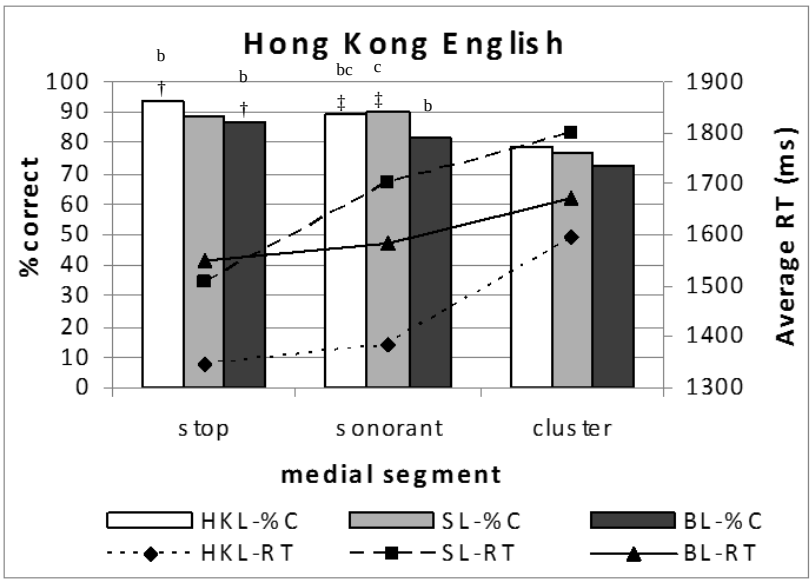
In addition to the overall patterns, we also looked at the difficulty posed by the different types of medial segments: stops, sonorants and clusters. Figure 2 gives the data for the three varieties.

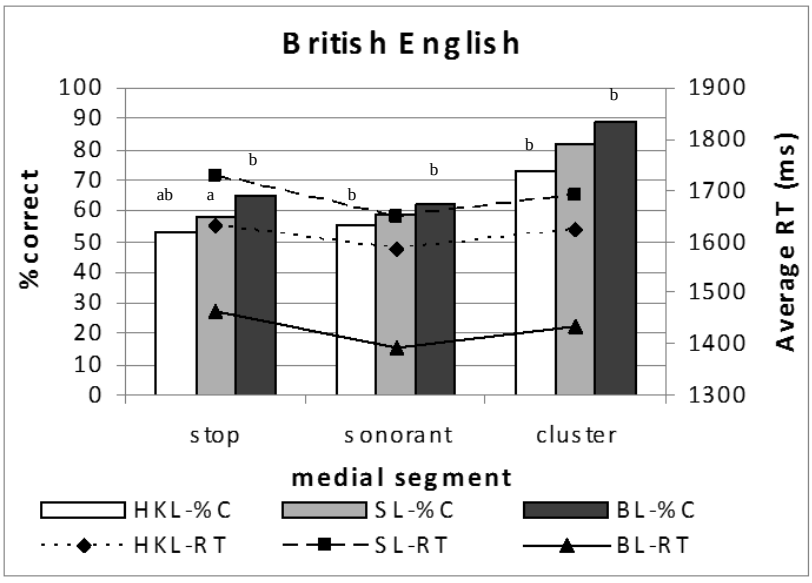
It is interesting to note that the patterns of difficulty are remarkably similar within each variety for each of the listener groups; in HKE it is easiest to discriminate pairs with medial stops, followed by pairs with sonorants, followed by pairs with clusters, whereas for BE the pattern is reversed (although there is little difference between the stops and sonorant pairs), and in SE sonorants are easiest, followed by clusters, then stops. Post hoc Bonferroni tests indicate the following:

- In HKE (Figure 2, top), Hong Kong listeners significantly outperform British listeners in terms of percentage correct on stops ($p<0.05$) and sonorant segments ($p<0.01$), and in RT on stops ($p<0.05$). However, Singaporean listeners do significantly better than Hong Kong listeners on sonorant segments in both percentage correct and RT (both $p<0.05$).
- Where BE is concerned (Figure 2, centre), for percentage correct, British listeners outperform Hong Kong listeners in word pairs containing stops ($p<0.001$), sonorants ($p<0.05$) and clusters ($p<0.001$), and outperform Singaporean listeners in stops ($p<0.01$); variations in reaction times are not significantly different.
- For SE (Figure 2, bottom), the only significant effect is in percentage correct for sonorants, where Singaporean listeners are significantly better than Hong Kong listeners ($p<0.01$). This is a rather puzzling statistic when one looks at the figure.

Statistical significance at $p \leq 0.05$ or greater is indicated on the figures as follows:

- %C: [a] = British listeners / Singapore listeners; [b] = British listeners /Hong Kong listeners; [c] = Singapore listeners / Hong Kong listeners.
- RT: † = British listeners /Hong Kong listeners; ‡ = Singapore listeners / Hong Kong listeners.





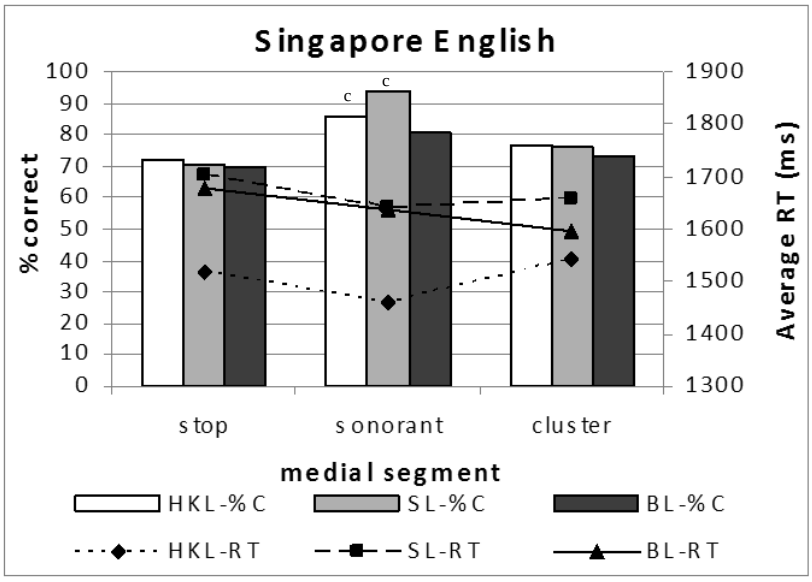


Figure 2: %C and RT for different medial consonants in each variety.

Discussion

We hypothesized that juncture boundaries in BE would be the most difficult for listeners in all three varieties to distinguish, and juncture boundaries in HKE would be the easiest. From these data it can be said that this is true. Where juncture cues are concerned, HKE appears to have the most obvious ones, as this is the variety in which all sets of listeners performed best. However, even though all listeners performed best on BE pairs containing clusters at the juncture, it is interesting to note that the British listeners perform a full 15 (average) percentage points better than their Hong Kong counterparts. We suggest that this may be because British listeners are more familiar with the subtle cues in BE clusters than the Hong Kong listeners, as clusters occur less frequently in HKE (Setter 2008), although there is also a chance that the fast reaction time indicates that the BE listeners are more complacent in their own variety and might have performed even better if they had not reacted so quickly. Nevertheless, the Hong Kong and Singaporean listeners still perform best on clusters in BE than on singleton stops or sonorants. It is likely that simply more juncture cues are contained in clusters than singleton pairs in BE than in the other varieties.

Singaporean listeners performed better on pairs with medial sonorant consonants in all three varieties. We do not have a suggestion as to why this might be the case; clearly, further investigation is needed. It is, however, interesting to find that different types of medial segments posed varying difficulty in juncture perception in the three varieties. It is possible that the listener groups performed consistently in terms of these patterns of medial segments across varieties because the cues for each segment type are consistent within each variety, or vary consistently in their level of perceptual difficulty. We have collected production data from groups of speakers of each variety and, once it is analysed, it will be useful to compare the perceptual differences with the production data to explore what contributes to such differences.

We also hypothesized that listeners would do best in accuracy and faster in RT in their own variety in comparison with the other listener groups. Concerning accuracy, there are two aspects to consider here: 1) whether the listener groups were more accurate than the other groups in their own variety, and 2) in which variety each listener group scored the highest. Where 1) is concerned, the Hong Kong and British listener groups both outperformed the other groups on HKE and BE respectively. However, in the case of 2), the British listeners scored highest overall on HKE. For all three listener groups, the reaction times were faster when listening to their own variety, although this is not statistically significant. However, the Singaporean listeners did best on HKE in terms of accuracy, and – more surprisingly – did not outperform the Hong Kong listeners on SE.

This leads us to conclude that, in terms of RT, familiarity with accent plays a part, as the listeners reacted more quickly to their own accent, if not (in the case of British and Singaporean listeners) more accurately. In addition, Hong Kong and Singaporean listeners outperformed the British listeners on both HKE and SE, which is probably because Hong Kong and Singaporean listeners are more familiar with accents with a strong Chinese influence.

Possibly the most reassuring result from the study so far is that all participants performed at above chance level in the identification of juncture pairs. This bodes well for international communication amongst speakers of these varieties of English as it indicates that there is generally enough information in the speech signal – even in RP – to make the processing of meaning successful among these three speaker groups.

Limitations

Our results show clear perception patterns among the three English varieties. There are, however, some limitations which need to be addressed in further studies.

Firstly, the stimuli were produced by only one speaker in each. We need to include

more speakers of each variety in future studies to be able to make generalizable comments about the results. HKE in particular is not a stable NVE and so it is dangerous to reach any firm conclusions based on this study. Nevertheless, it is worth pointing out that our results based on one 'typical' speaker concur very well with other independent studies involving multiple speakers (Kirkpatrick and Saunders 2005; Kirkpatrick et al. 2008), and we believe that our results will be reliably replicated with more speakers.

Secondly, the stimuli were recorded in a formal setting, i.e., they were not taken from natural conversations, and it is therefore possible that the speakers hyper-articulated their speech to some extent (see, e.g., Lindblom 1990), even though the word pairs were placed in an unstressed position in the carrier phrase. It is possible that different – or, even, less successful – perception patterns may emerge with more naturalistic materials, and it would also be useful to replicate the study in a communicative context. It is worth mentioning in this connection that we did not take speaking rate into account or control for it; while this may have some bearing on the results presented here, it would be more difficult to control for speaking rate were naturalistic stimuli to be used.

Finally, we need evidence from acoustic analysis – including speaking rate – to be able to comment more confidently on the reasons for the results we have presented.

Conclusion

This study is the first one to investigate the perception of phonetic juncture cues across OVEs and NVEs, and so makes an important contribution to the research in this area. We have provided clear, controlled experimental data on the perception of the three varieties across the three listener groups. Studies such as Kirkpatrick (1995) and Kirkpatrick et al. (2008) use stimulus data which is not controlled for content and which therefore make it difficult to comment on the precise issues surrounding intelligibility.

We aim to expand the study of intelligibility in NVEs and other varieties of English,

including learner varieties, to include speakers from continental Europe and other parts of the Southeast Asia region. It will be interesting to see what issues arise when more speakers of English are included in the study and how HKE compares with other varieties in terms of intelligibility. In future research, as mentioned above, we also intend to acoustically examine the phonetic features of the juncture cues in speakers from each variety in order to pinpoint the exact nature of the acoustic cues produced by speakers of different varieties of English.

Acknowledgements

This study benefitted from a British Council / Association of Commonwealth Universities' grant for international collaboration (project number CC090034) awarded to the first and second authors, and also from a Direct Grant for Research 2009-10, First Round from the Chinese University of Hong Kong awarded to the second author. We would like to thank Mr Chris Ryder for his help collecting the UK data for this study.

References

Brown, Adam (2005) A bibliography on Singapore English pronunciation. In David Deterding, Adam Brown and Ee Ling Low (eds.) *English in Singapore: Phonetic Research on a Corpus* (pp. 184 – 202). Singapore: McGraw Hill Education.

Boucher, V. J. (1988) A parameter of syllabification for VstopV and relative-timing invariance. *Journal of Phonetics* 16, 299-326.

Chan, Alice Y. W. and Li, David C. S. (2000) English and Cantonese phonology in contrast: explaining Cantonese ESL learners' English pronunciation problems. *Language, Culture and Curriculum* 13/1: 67-85.

Christie, W. M. 1974. Some cues for syllable juncture perception in English. *Journal of the Acoustical Society of America* 55, 819-821.

Deterding, David and Low, Ee Ling (2001) The NIE corpus of spoken Singapore English (NIECSSE). *SAAL Quarterly* 56, 2-5.

Deterding, David, Wong, Jennie and Kirkpatrick, Andy (2008) The pronunciation of Hong Kong English. *English World-Wide* 29, 148-175.

Gårding, Eva (1967) *Internal Juncture in Swedish.* Travaux de l'Institut de phonétique de Lund 6. Lund, Sweden: CWK Gleerups.

Gargesh, R. (2008). Indian English: phonology. In Rajend Mesthrie (ed.) *Varieties of English 4: Africa, South and Southeast Asia* (pp. 231 – 243). Berlin and New York: Mouton de Gruyter.

Hardcastle, William J. and Marchal, Alain (1990) *Speech Production and Speech Modelling*. Berlin: Springer.

Hawkins, Sarah (2003) Roles and representations of systematic fine phonetic detail in speech understanding. *Journal of Phonetics* 31, 373-405.

Hung, Tony (2000) Towards a phonology of Hong Kong English. *World Englishes* 19/3, 337-356.

Hung, Tony (2005) Word stress in Hong Kong English: a preliminary study. *Hong Kong Baptist University Papers in Applied Language Studies* 9, 29-40.

James, Allan. and Leather, Jonathan (2002) *New Sounds 2000: Proceedings of the Fourth International Symposium on the Acquisition of Second-Language Speech.* Klagenfurt, Austria: University of Klagenfurt.

Kim, Dahee, Stephens, Joseph D. W. and Pitt, Mark A. (2012) How does content play a part in splitting words apart? Production and perception of word boundary cues in casual speech. *Journal of Memory and Language* 66, 509-529.

Lehiste, Ilse (1960) An acoustic-phonetic study of internal open juncture. *Phonetica* 5, 1-54.

Kirkpatrick, Andy, Deterding, David and Wong, Jennie (2008) The international intelligibility of Hong Kong English. *World Englishes* 27/3-4, 359-377.

Kirkpatrick, Andy and Saunders, Neville (2005) The intelligibility of Singaporean English: a

case study in an Australian university. In David Deterding, Adam Brown and Ee Ling Low (eds.) *English in Singapore: Phonetic Research on a Corpus* (pp. 153 – 162). Singapore: McGraw Hill Education.

Lindblom, B. (1990) Explaining phonetic variation: a sketch of the H&H theory. In William J. Hardcastle and Alain Marchal (eds.) *Speech Production and Speech Modelling* (pp. 403 – 431). Berlin: Springer.

Low, Ee Ling and Azirah, Hashim (2012) *English in Southeast Asia: Features, Policy and Language in Use*. Amsterdam and Philadelphia, PA: John Benjamins B. V.

Luke, Kang Kwong (2008) Stress and intonation in Hong Kong English. Paper given at *The 14th International Association of World Englishes Conference*, City University, Hong Kong, December 2008.

Luke, Kang Kwong and Richards, Jack (1982) English in Hong Kong: functions and status. *English World-Wide* 3, 47-64.

Li, David C. S. (1999) The functions and status of English in Hong Kong: A post-1997 update. *English World-Wide* 20/1, 61–110.

Mattys, Sven L. and Melhorn, James F. (2007) Sentential, lexical, and acoustic effects on the perception of word boundaries. *Journal of the Acoustical Society of America* 122, 554-567.

Nakatani, Lloyd H. and Dukes, Kathleen D. (1977) Locus of segmental cues for word juncture. *Journal of the Acoustical Society of America* 62, 714-719.

Quené, Hugo (1993) Segment durations and accent as cues to word segmentation in Dutch. *Journal of the Acoustical Society of America* 94, 2027-2035.

Redford, Melissa A. and Randall, Patrick (2005) The role of juncture cues and phonological knowledge in English syllabification judgements. *Journal of Phonetics* 33, 27-46.

Rietveld, A. C. M. (1980) Word boundaries in the French language. *Language and Speech* 23,

289-296.

Schneider, Edgar (2007) *Postcolonial English: Varieties Around the World*. Cambridge: Cambridge University Press.

Schwab, Sandra, Miller, Joanne L., Grosjean, Françoise and Mondini, Michèle (2008) Effect of speaking rate on the identification of word boundaries. *Phonetica* 65, 173-186.

Setter, Jane (2006) Speech rhythm in world Englishes: The case of Hong Kong. *TESOL Quarterly* 40/4, 763-782.

Setter, Jane (2008) Consonant clusters in Hong Kong English. *World Englishes* 27/3-4, 502-515.

Setter, Jane, Wong, Cathy S. P. and Chan, Brian H. S. (2010) *Hong Kong English*. Edinburgh: Edinburgh University Press.

Tayao, Ma. Lourdes G. (2008) Philippine English: phonology. In Rajend Mesthrie (ed.) *Varieties of English 4: Africa, South and Southeast Asia* (pp. 292 – 306). Berlin and New York: Mouton de Gruyter.

Trask, R. L. (1996) *A Dictionary of Phonetics and Phonology*. London: Routledge.

Treiman, Rebecca and Danis, Catalina (1988) Syllabification of intervocalic consonants. *Journal of Memory and Language* 27, 87-104.

Treiman, Rebecca and Zukowski, Andrea (1990) Toward an understanding of English syllabification. *Journal of Memory and Language* 29, 66-85.

Wong, Cathy S. P. (2004) Does Cantonese lexical tone affect the acquisition of English word stress? *Proceedings of the Applied Linguistics Association of Korea's 2004 Annual International Conference and General Meeting* (pp. 193-198). Korea: Hanyang University, Korea.

Wong, Cathy S. P. (1991) The stress patterns of nonsense English words of Cantonese-speaking ESL learners. *Chinese University of Hong Kong Papers in Linguistics* 3, 83-111.

Wong, Cathy S. P. and Setter, Jane (2002) Is it 'night' or 'light'? How and why Cantonese-speaking ESL learners confuse syllable-initial [n] and [l]. In Allan James and Jonathan Leather (eds.) *New Sounds 2000: Proceedings of the Fourth International Symposium on the Acquisition of Second-Language Speech* (pp. 351 - 359). Klagenfurt, Austria: University of Klagenfurt.

APPENDIX

Consonant	**Juncture before consonant**	**Juncture after consonant**
Single obstruents		
/p/ + vowel	why pink	wipe ink
	grey pale	grape ale
	key part	keep art
/t/ + vowel	my take	might ache
	grey ties	great eyes
	knee tape	neat ape
/k/ + vowel	buy coil	bike oil
	may coat	make oat
	we cash	weak ash
/d/ + vowel	free Danny	freed Annie
	gray day	Grade A
	stay dill	stayed ill
/tʃ/	lawn chair	launch air
	why choose	white shoes
Single sonorants		
/l/ + vowel	see lying	seal eyeing
	we loan	we'll own
	die lies	dial eyes
/n/ + vowel	no notion	known ocean
/m/ + vowel	hoe maker	home acre
	clay mice	claim ice
Consonant clusters		
/t/ + /r/	my train	might rain
	buy trade	bite raid
/s/ + /t/	keep sticking	keeps ticking
/s/ + /pr/	it sprays	it's praise

Table 1. List of juncture pairs used in the main study

Variety												
Listener Group	BE				SE				HKE			
	Average	Min	Max	St. Dev	Average	Min	Max	St. Dev	Average	Min	Max	St. Dev
BL	71.46[b]	51	79	6.22	74.17[c]	52	94	8.11	81.71[bc]	51	95	9.28
HKL	59.58[ab]	51	69	4.01	78.33[ac]	73	83	3.35	89.50[bc]	80	97	3.83
SL	64.48[ab]	54	69	4.15	78.33[ac]	66	84	5.57	86.35[bc]	72	90	5.32

Table 2. %C) calculated to two decimal points. Statistical significance at $p \leq 0.05$ for means: [a] = BE/SE; [b] = BE/HKE; [c] = SE/HKE.

Variety												
Listener Group	BE				SE				HKE			
	Average	Min	Max	St. Dev	Average	Min	Max	St. Dev	Average	Min	Max	St. Dev
BL	1438.95	1113	1896	225.51	1637.66	1127	2411	366.37	1570.78	1163	2237	271.20
HKL	1582.04	896	2431	407.97	1472.72	982	2410	337.40	1387.09	979	2246	302.12
SL	1605.31	1204	2171	289.04	1561.02	1175	19.5	315.09	1570.02	1286	1931	195.40

Table 3. RT in milliseconds calculated to two decimal points. No statistical significance.