

Phonetic correlates of unintelligibility in Vietnamese-accented English

Una Cunningham

School of Arts and Media, Dalarna University, Falun

Abstract

Vietnamese speakers of English are often able to communicate much more efficiently in writing than in speaking. Many have quite high proficiency levels, with full command of advanced vocabulary and complex syntax, yet they have great difficulty making themselves understood when speaking English to both native and non-native speakers. This paper explores the phonetic events associated with breakdowns in intelligibility, and looks at compensatory mechanisms which are used.

Intelligibility

The scientific study of intelligibility has passed through a number of phases. Two strands that have shifted in their relative prominence are the matter of to whom non-native speakers are to be intelligible. In one strand the emphasis is on the intelligibility of non-native speakers to native English-speaking listeners (Flege, Munro et al. 1995; Munro and Derwing 1995; Tajima, Port et al. 1997). This was the context in which English was taught and learned – the majority of these studies have been carried out in what are known as inner circle countries, which, in turn, reflects the anglocentricism which has characterised much of linguistics. The other strand focuses on the position of English as a language of international communication, where intelligibility is a two-way affair between a native or non-native English-speaking speaker and a native or non-native English-speaking listener (Irvine 1977; Flege, MacKay et al. 1999; Kirkpatrick, Deterding et al. 2008; Rooy 2009).

The current study is a part of a larger study of how native speakers of American English, Swedish, Vietnamese, Urdu and Ibo are perceived by listeners from these and other language backgrounds. Vietnamese-accented speech in English has been informally observed to be notably unintelligible for native English-speaking listeners and even for Vietnamese listeners there is great difficulty in choosing which of four words has been uttered (Cunningham 2009).

There are a number of possible ways in which intelligibility can be measured. Listeners can be asked to transcribe what they hear or to choose from a number of alternatives. Stimuli vary from spontaneous speech through texts to sentences to wordlists. Sentences with varying degrees of semantic meaning are often used (Kirkpatrick, Deterding et al. 2008) to control for the effect of contextual information on intelligibility. Few intelligibility studies appear to be concerned with the stimulus material. The question of what makes an utterance unintelligible is not addressed in these studies. The current paper is an effort to come some way to examining this issue.

Learning English in Vietnam

The pronunciation of English presents severe challenges to Vietnamese-speaking learners. Not only is the sound system of Vietnamese very different from that of English, but there are also extremely limited opportunities for hearing and speaking English in Vietnam. In addition, there are limited resources available to teachers of English in Vietnam so teachers are likely to pass on their own English pronunciation to their students.

University students of English are introduced to native-speaker models of English pronunciation, notably Southern educated British, but they do not often have the opportunity to speak with non-Vietnamese speakers of English. Most studies of Vietnamese accents in English have been based in countries where English is a community language, such as the U.S. (Tang 2007) or Australia (Nguyen 1970; In-gram and Nguyen 2007). This study is thus unusual in considering the English pronunciation of learners who live in Vietnam. The speech material presented here was produced by members of a group of female students from Hanoi.

Vietnamese accents of English

The most striking feature of Vietnamese-accented English is the elision of consonants, in particular in the syllable coda. This can obvi-

ously be related to the phonotactic constraints operational in Vietnamese, and it is clearly a problem when speaking English which places a heavy semantic load on the coda in verb forms and other suffixes. Consonant clusters are generally simplified in Vietnamese-accent English to a degree that is not compatible with intelligibility. Even single final consonants are often absent or substituted for by another consonant which is permitted in the coda in Vietnamese.

Other difficulties in the intelligibility of Vietnamese-accented English are centred in vowel quality. English has a typologically relatively rich vowel inventory, and this creates problems for learners with many L1s, including Vietnamese. The distinction between the vowels of KIT and FLEECE to use the word class terminology developed by John Wells (Wells 1982) or *ship* and *sheep* to allude to the popular pronunciation teaching textbook (Baker 2006) is particularly problematic.

Other problematic vowel contrasts are that between NURSE and THOUGHT (e.g. *work* vs *walk*) and between TRAP AND DRESS (e.g. *bag* vs *beg*). The failure to perceive or produce these vowel distinctions is a major hinder to the intelligibility of Vietnamese-accented English.

Vowel length is not linguistically significant in Vietnamese and the failure to notice or produce pre-fortis clipping is another source of unintelligibility. Another interesting effect that is attributable to transfer from Vietnamese is the use in English of the rising *sac* tone on syllables that have a voiceless stop in the coda. This can result in a pitch prominence that may be interpreted as stress by listeners.

Vietnamese words are said to be generally monosyllabic, and are certainly written as monosyllables with a space between each syllable. This impression is augmented (or possibly explained) by the apparent paucity of connected speech phenomena in Vietnamese and consequently in Vietnamese-accented English.

Analysis

A number of features of Vietnamese-accented English will be analysed here. They are a) the vowel quality distinction between the words *sheep* and *ship*, b) the vowel duration distinction between *seat* and *seed*, and c) the causes of global unintelligibility in semantically meaningful sentences taken from an earlier study (Munro and Derwing 1995).

Vowel quality

The semantic load of the distinction between the KIT and FLEECE vowels is significant. This opposition seems to be observed in most varieties of English, and it is one that has been identified as essential for learners of English to master (Jenkins 2002). Nonetheless, this distinction is not very frequent in the languages of the world. Consequently, like any kind of new distinction, a degree of effort and practice is required before learners with many first languages, including Vietnamese, can reliably perceive and produce this distinction.

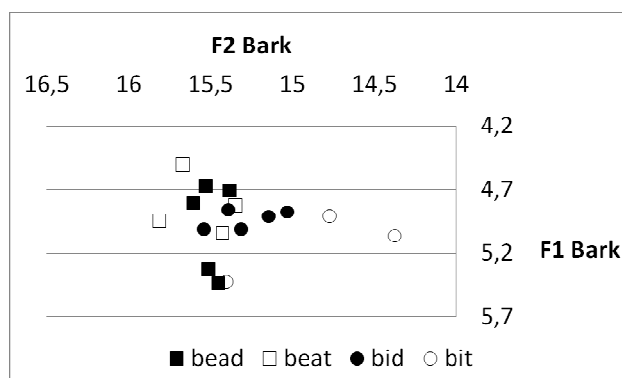


Figure 1. F1 vs F2 in Bark for S15 for the words *bead*, *beat*, *bid*, *bit*.

Fig.1 shows the relationship between F1 and F2 in Bark for the vowels in the words *beat*, *bead*, *bit* and *bid* for S15, a speaker of Vietnamese (a 3rd year undergraduate English major student at a university in Hanoi). This speaker does not make a clear spectral distinction between the vowels. As spectral quality is the most salient cue to this distinction for native speakers of English (Ladefoged 2006; Cruttenden 2008), the failure to make a distinction is obviously a hinder to intelligibility.

Vowel duration

Enhanced pre-fortis clipping is used in many varieties of English as a primary cue to post-vocalic voicelessness (Ladefoged 2006; Cruttenden 2008). It has been well documented that phonologically voiced (lenis) post-vocalic consonants are often devoiced by native speakers of many varieties of English (e.g. Cruttenden 2008). This means that in word sets such as *bead*, *beat*, *bid*, *bit* native speakers will signal postvocalic voicing primarily by shortening the prevocalic vowel in *beat* and *bit*. In addition, native speakers will have a secondary durational cue to the *bead*, *beat* vs *bid*, *bit* vowel distinction where the former vowel is system-

atically longer than the latter (Cruttenden 2008).

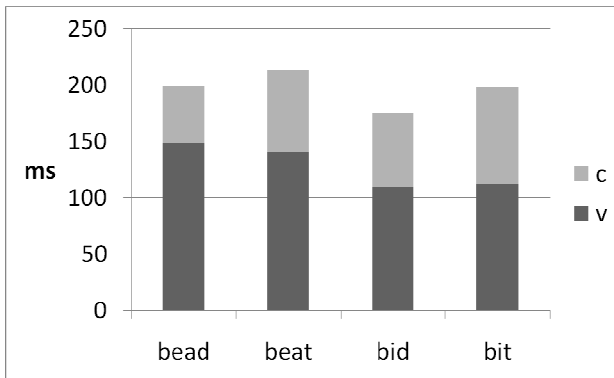


Figure 2. Average vowel and stop duration in ms for S15 for 5 instances of the words *bead*, *beat*, *bid*, *bit*.

So, as is apparent from Figure 2, speaker S15 manages to produce somewhat shorter vowels in *bid* and *bit* than in *beat* and *bit*. This is the primary cue that this speaker is using to dissimilate the vowels, although not, unfortunately the cue expected as most salient by native listeners. But there is no pre-fortis clipping apparent. This important cue to post-vocalic voicing is not produced by this speaker. In conjunction with the lack of spectral distinction between the vowels of *bead*, *bead* vs. *bid*, *bit* seen in Figure 1, the result is that these four words are perceived as indistinguishable by native and not-native listeners (Cunningham 2009).

Sentence intelligibility

A number of factors work together to confound the listener of Vietnamese-accented English in connected speech. Not only is it difficult to perceive vowel identity and post vocalic voicing as in the above examples, but there are a number of other problems. Consider the sentence *My friend's sheep is often green*. This is taken from the stimuli set used for the larger study mentioned above. The advantage of sentences of this type is that the contextual aids to interpretability are minimised while connected speech phenomena are likely to be elicited. There are here a number of potential pitfalls for the Vietnamese speaker of English. The cluster at the end of *friend's*, especially in connection with the initial consonant in *sheep* can be expected to prove difficult and to be simplified in some way. The quality and duration of the vowels in *sheep* and *green* can be expected to cause con-

fusion (as illustrated in Figures 1 and 2 above. The word *often* is liable to be pronounced with a substitution of /p/ for the /f/ at the end of the first syllable, as voiceless stops are permissible in coda position in Vietnamese while fricatives are not.

Let us then see what happens when speaker V1, a 23-year old male graduate student of English from Hanoi, reads this sentence. In fact he manages the beginning of the utterance well, with an appropriate (native-like) elision of the /d/ in *friends*. Things start to go wrong after that with the word *sheep*. Figure 3 shows a spectrogram of this word using Praat (Boersma and Weenink 2009). As can be seen, the final consonant comes out as an ungrooved [s].

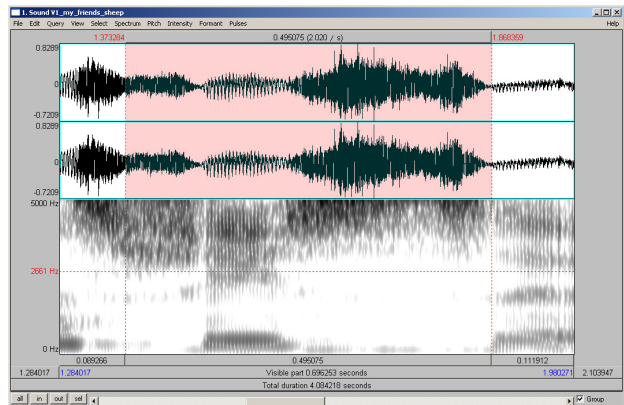


Figure 3. The word *sheep* as pronounced by speaker V1.

Now there is a possible explanation for this. As mentioned above, final /f/ as in the word *if* is often pronounced as [ip]. This pronunciation is viewed as characteristic of Vietnamese-accented English in Vietnam – teacher and thus learner awareness of this is high, and the feature is stigmatised. Thus the avoidance of the final /p/ of *sheep* may be an instance of hypercorrection. It is clearly detrimental to V1's intelligibility.

Another problematic part of this sentence by V1 is that he elides the /z/ in the word *is*. There is silence on either side of this vowel. Again, this leads to intelligibility problems. The final difficulty for the listener in this utterance is a matter of VOT in the word *green*. V1 has no voice before the vowel begins as is shown in figure 4. The stop is apparently voiceless and the release is followed by a 112ms voiceless aspiration. This leads to the word being misinterpreted by listeners as *cream*.

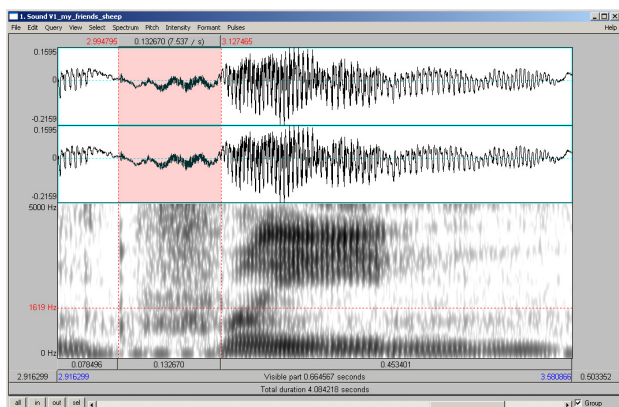


Figure 4 the word *green* as pronounced by speaker V1. The marking shows aspiration after the release of the initial stop.

Conclusion

So it can be seen that the intelligibility of these Vietnamese speakers of English is a major problem for them and their interlocutors. Not only do they have non-native pronunciation features that are clear instances of transfer from their L1, Vietnamese, they also have other, spontaneous, modifications of the target sounds. This is part of the general variability that characterises non-native pronunciation, but when the sounds produced are as far from the target sounds as they are in the speech of V1, communication is an extreme effort.

References

- Baker, A. (2006). *Ship or sheep: an intermediate pronunciation course*. Cambridge, Cambridge University Press.
- Boersma, P. and D. Weenink (2009). *Praat: doing phonetics by computer*.
- Cruttenden, A. (2008). *Gimson's Pronunciation of English*. London, Hodder Arnold
- Cunningham, U. (2009). Quality, quantity and intelligibility of vowels in Vietnamese-accented English. *Issues in Accents of English II: Variability and Norm*. E. Waniek-Klimczak. Newcastle, Cambridge Scholars Publishing Ltd.
- Flege, J. E., I. R. A. MacKay, et al. (1999). Native Italian speakers' perception and production of English vowels. *Journal of the Acoustical Society of America* 106(5): 2973-2987.
- Flege, J. E., M. J. Munro, et al. (1995). Factors affecting strength of perceived foreign accent in a 2nd language. *Journal of the Acoustical Society of America* 97(5): 3125-3134.
- Ingram, J. C. L. and T. T. A. Nguyen (2007). Vietnamese accented English: Foreign accent and intelligibility judgement by listeners of different language backgrounds, University of Queensland.
- Irvine, D. H. (1977). Intelligibility of English speech to non-native English speakers. *Language and Speech* 20: 308-316.
- Jenkins, J. (2002). A sociolinguistically based, empirically researched pronunciation syllabus for English as an international language. *Applied Linguistics* 23(1): 83-103.
- Kirkpatrick, A., D. Deterding, et al. (2008). The international intelligibility of Hong Kong English. *World Englishes* 27(3-4): 359-377.
- Ladefoged, P. (2006). *A Course in Phonetics*. Boston, Mass, Thomson.
- Munro, M. J. and T. M. Derwing (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech* 38: 289-306.
- Nguyen, D. L. (1970). *A contrastive phonological analysis of English and Vietnamese*. Canberra, Australian National University.
- Rooy, S. C. V. (2009). Intelligibility and perceptions of English proficiency. *World Englishes* 28(1): 15-34.
- Tajima, K., R. Port, et al. (1997). *Effects of temporal correction on intelligibility of foreign-accented English*, Academic Press Ltd.
- Tang, G. M. (2007). Cross-linguistic analysis of Vietnamese and English with implications for Vietnamese language acquisition and maintenance in the United States. *Journal of Southeast Asian American Education and Advancement* 2.
- Wells, J. C. (1982). *Accents of English*. Cambridge, Cambridge University Press