MUTUAL INTELLIGIBILITY OF MALAY- AND SWEDISH-ACCENTED ENGLISH: AN EXPERIMENTAL STUDY

Hyeseung Jeong[1]
Bosse Thorén[2]
Juliana Othman[3]
Department of Language and Literacy Education at the Faculty of Education, University of Malaya[1,3]
Dalarna University, Sweden[2]
hjeong@um.edu.my

First received: 3 November 2016 Final proof received: 30 May 2017

Abstract

In using English as an international language (EIL), one important issue is mutual intelligibility among EIL speakers from different language backgrounds. The present study investigates the cross-linguistic intelligibility of Malay-accented English and Swedish-accented English, regarding the three phonetic features – word stress pattern, consonant clusters, and long vowel in particular. We prepared 15 English statements that are evidently true or false if understood, and examined to what extent the three phonetic features are related to 30 Swedish and 38 Malaysian listeners' understandings of the statements read by a speaker from the other language group. We compared the Malaysian and Swedish listeners' answers given with understanding as well as processing time to respond. The listeners' own accounts of their struggles in understanding the speakers' pronunciations were also analyzed. Results show that Malaysian listeners easily understood Swedish-accented English, while Swedish listeners struggled to understand Malay-accented English. The difference between the two groups of listeners seems to be closely related to the degree of the realization of the three phonetic features by the speakers as well as to the degree of the use of these features as perceptual cues by the listeners. Based on the findings, we discuss potential phonetic core features of EIL for intelligibility and some pedagogical implications for teaching English pronunciation to the learners of the language.

Keywords: EIL; intelligibility of EIL pronunciation; phonetic core features; Malay-accented English; Swedish-accented English

Intelligibility refers to the "degree of match between a speaker's intended message and the listener's comprehension" (Derwing & Munro, 2015, p. 5). The concept therefore provides useful ways to measure to what extent the speaker's pronunciation is understood by the listener.

In the context of using English as an international language (EIL), intelligibility has been suggested to be the alternative to nativelikeness or nativelike accent for the norm or ideal goal of second language (L2) pronunciation (Derwing & Munro, 2015; Jenkins, 2002; Levis, 2005; Rajadurai, 2006; Sewell, 2016; Thomson, 2014). The resetting of the criteria for assessing an L2 pronunciation of English from nativelikeness to intelligibility involves the idea that it is not essential for L2 speakers to change their own accent to sound nativelike. However, it is also not for arguing that any L2 pronunciation is unproblematic. Researchers advocating the intelligibility principle suggest that there are phonetic features crucial for intelligible speech, and EIL speakers need to strive to realize these features properly in order to be well understood by their interlocutors (e.g., Derwing & Munro, 2015; Hahn, 2004; Jenkins, 2002; Levis, 2005; Quené & van Delft, 2010; Rajadurai, 2006; Zielinski, 2008).

For the emergence of the intelligibility principle in the context of EIL, several factors can be considered. First of all, the population of English non-native speakers (NNS) greatly outnumber that of English native speakers (NS), resulting in numerously different English accents (McKay, 2009). Even among English NSs, there is a wide variety of pronunciations or accents that can considerably differ from one another (see Best, 2016). In such a situation, using nativelikeness as the index for assessing the English speaker's pronunciation has involved the issue of construct validity. In addition, some features of English native speakers' pronunciation, such as elisions or assimilations, can be rather difficult for non-native speakers of the language (Jenkins, 2002). Sounding English nativelike therefore does not automatically guarantee a person's pronunciation to be perceived intelligible. Moreover, it has been found that the correlation between the perceived degree of foreign accent and intelligibility is partial and weak (Munro & Derwing, 2013).

The intelligibility principle has been supported increasingly by L2 pronunciation researchers. While

 doi: dx.doi.org/10.17509/ijal.v7i1.6857

aligning with these researchers, we perceive two issues in previous studies carried out under this principle. Firstly, most of the previous studies evaluated the intelligibility of L2 English pronunciations only against the perceptions of native English listeners (e.g., Hahn, 2004; Munro & Derwing, 1995; Quené & van Delft, 2010; Zielinski, 2008). Regarding this, earlier Jenkins (2002), Kashiwagi and Snyder (2008), and more recently Sewell (2016) argued that L2 pronunciation should be assessed not only against the perceptions of English NSs but also against the perceptions of NNSs from different language backgrounds, given the global status of English. Secondly, the findings of previous studies on whether and to what extent certain phonetic features influence the intelligibility of L2 pronunciation are inconsistent and inconclusive. In the following paragraphs, we elaborate this second issue, in relation to three phonetic features that we focus on in the present study, namely word stress patterns, consonant clusters, and vowel length.

First, based on English NSs' perception, several studies identified stress pattern as crucial for intelligible L2 speech (e.g., Hahn, 2004; Rajadurai, 2006; Tiffen, 1992; Zielinski, 2008). While acknowledging that stress pattern can be an important perceptual cue for NSs, Jenkins (2002) and Kashiwagi, Snyder, and Craig (2006) argued that it is not a core feature that affects NNS-NNS communications seriously. However, Field (2005), and Kashiwagi and Snyder (2008) found out that stress pattern had a powerful effect on the intelligibility of L2 English pronunciation when L2 Japanese speakers were assessed against both American (NS) and Japanese judges (NNS). In addition, the data of Jenkins (2002) and Kashiwagi et al. (2006) do not include L2 speakers of English whose L1 depends heavily on stress pattern as perceptual cue, such as the Swedish (Abelin & Thorén, 2015).

Second, consonant clusters in general have been suggested as what affect the intelligibility of L2 pronunciation negatively when they are not clearly realized (Jenkins, 2002; Lesley, 2014; Suenobu, Kanzaki, & Yamane, 1992; Tajima, Port, & Dalby, 1996; Zhang, 2015). However, Sewell (2016), in his study on Hong Kong English, limited the importance of articulating consonant clusters to the pronunciation of mono syllabic words (e.g., [-rd/-d] in card, or [-st] in post), or to the cases where consonant clusters carry grammatical information (e.g., [-d] or [-t] for the past tense morpheme *-ed*). He argued that consonant reduction does not hurt intelligibility of polysyllabic words (e.g., *department*) because the listener would get sufficient phonetic information for recognizing these words even if not hearing all consonant sounds in them.

Third, some studies reported that vowel length or durational realization of the English tense-lax contrast is a factor that possibly affects international communications, based on empirical data (Bent, Bradlow, & Smith, 2008; Jenkins, 2002; Quené & van Delft, 2010; Rajadurai, 2006; Smith et al., 2003; Tajima et al., 1996) For example, according to Jenkins (2002) keeping contrast between long and short vowel, mostly /iː/ - /ɪ/ as in *leave – live*, is a phonetic core feature of intelligible speech in using English in international contexts. By contrast, Munro and Derwing (2015) argue that the phonetic contrast in the same pairs of words that affects the intelligibility of speech is not durational but it is more of quality difference between the contrasting vowels.

Considering the mixed results about the effects of the three phonetic features on intelligible speech in the previous studies, this study aims to further examine to what extent these features affect cross-linguistic communication. We also seek to use the perceptions of English NNSs rather than those of English NSs in view of scarcity of studies that test L2 pronunciation against L2 listeners. For these research objectives, we looked for two L2 groups, one of which tends to reduce sounds in consonant clusters and not to use stress placement and vowel length as ways to cue the meanings of words, while the other has tendency to rely on the clear realization of the three phonetic features for both delivering their intended speech and understanding others' utterances. We consider that comparing the mutual perceptions of two such groups on each other's English pronunciation would amplify the relationship between the realization of the three phonetic features and the intelligibility of L2 speech

The extant literature informs us that Malaysian and Swedish users of English would be suitable participant groups for our research purpose (Abelin & Thorén, 2015, 2016; Brown, 1988; Davidsen-Nielsen & Harder, 2001; Elert, 1964; Garlén, 1988; Phoon, Abdullah, & Maclagan, 2013; Yong, 2001).

Firstly, Malaysian English speakers tend to stress the final syllable of the final word in a phrase, altering commonly known word stress patterns more or less (Brown, 1988; Yong, 2001). They also often omit sounds in consonant clusters or insert epenthesis vowels, due to the influence of Malay language in which most words have consonant-vowel (CV) and consonant-vowel-consonant (CVC) sequence in syllables (Yong, 2001). As for the tense-lax vowel contrast in Malaysian English, there is disagreement among researchers. Pillai, Mohd. Don, Knowles, and Tang (2010) report that Malaysian English in general lacks contrast in vowel quality, but it has contrast in vowel duration. By contrast, Baskaran (2008) and Yong (2001) described that, in Malaysian English, the vowel sounds of minimal word pairs (e.g., *bit/beat*,

pill/peal/, full/fool, cot/caught) are pronounced with similar vowel duration.

On the other hand, according to Abelin and Thorén (2015), native Swedish speakers centrally use word stress placement to cue the meanings of the words. When listening to others, they also seldom recognize words if the stress pattern of these words is altered. In addition, since Swedish and English have very similar phonotactic rules that allow multiple consonants in a row (up to 5, and usually 2 or 3) (Garlén, 1988), Swedish speakers would pronounce English consonant clusters without omitting sounds, and also may expect similar pronunciation patterns from their interlocutors. Moreover, one can easily expect a clear long-short vowel contrast in Swedish English, given that Swedish has a quantity contrast based mainly on the durations of vowel and subsequent consonant in stressed syllables (V:C - VC:) (Elert, 1964; Thorén, 2003), and also displaying quality differences between long-short vowel allophones.

Reflecting on the literature and the objectives of the study, we formulated the following research questions:

1. Which linguistic group – Malaysian or Swedish users of English – struggle more to understand the English pronunciation of the speaker from the other group?
2. Do the contrasting phonetic features between Malay-accented English and Swedish-accented English—stress pattern, consonant clusters, and vowel length—negatively affect the intelligibility of the two varieties of English?

METHOD

Speakers

Speakers were one Malay L1 speaker and one Swedish L1 speaker. The Malaysian speaker was a 20-year-old female undergraduate at the English teacher training program at a Malaysian university. She learned English in Malaysia. The Swedish speaker was a 63-year-old male working at a Swedish university as a lecturer. He grew up and learned English in the Stockholm area. Both speakers had experience of interacting with various English speakers from different language backgrounds. They were assessed as having a moderate Malay accent and Swedish accent, respectively, by four international people experienced with both Malay-accented English and Swedish-accented English.

Stimulus sentences

Fifteen English statement sentences consisting of 5 to 11 words (M=7.07) were created for the stimulus material, which both of the Malaysian and Swedish speakers read. The statements were assumed to be easily determined as true or false when they are understood, like "A trumpet is a musical instrument". Word choices for the sentences were made in a way to test three phonetic variables, namely words containing 2-5 syllables to test stress placement, words containing consonant clusters, and sentences with 2-4 words containing a long vowel. That is, among the 15 sentences, 5 sentences had polysyllabic words, 5 sentences had consonant cluster words, and 5 sentences had long vowel words. In the course of checking with Malaysian and Swedish listener participants which words they did not hear clearly (see the procedure section), we also asked them if there were words that they did not know the meaning of. One Malaysian participant said he knew the word *fists* in a sentence, but did not use it frequently, and one Swedish participant said she had not known the word before. Another Swedish participant did not know the meaning of the word *pond* in another sentence. Except for these cases, listener participants understood the meanings of the sentences easily and quickly while they were reading them on a paper.

Recordings and material preparation

The recording of the Malaysian and Swedish speakers was done at a university music studio by a technician. The two speakers were asked to read the fifteen stimulus sentences several times as naturally as possible with their own accents. All the readings were recorded at 48,000 Hertz sampling frequency to ensure good sound quality. Then, the recordings were saved as Microsoft WAVE files. Beside the 15 stimulus sentences, a sample sentence was read by the same speakers and audio-recorded for the purpose of checking sound volume before testing each listener. After audio-recording, both speakers' sound files were edited with the software Praat (Boersma & Weenink, 2014) to be ready for the experiment. 15 items (sentences) selected from several versions were randomly organized with 7-second silence intervals between them. We created two versions of each speaker's sound file, with reversed order of the sentences in order to eliminate possible training effects during the experiment.

Listeners

38 Malaysian and 30 Swedish listeners were recruited through convenience sampling. Malaysian listeners were either staff or students at a Malaysian university, whose ages ranged from 19 to 75. Swedish listeners were either staff or students at two Swedish universities, from 25 to 67 years old. All the participants began to learn English at the ages of 2 to 13, and were using English on a daily or regular basis. We did not give participants a hearing test before the actual experiment, but all of them voluntarily participated, clearly knowing that they were going to listen to English sentences, and none of them reported hearing problems that may affect their listening performance. The Malaysian participants were paid 5 Malaysian Ringgits, and the

Swedish participants were given a Malaysian souvenir key chain for participation.

Data collection procedure

Each participant was tested individually in a quiet room for 10 to 15 minutes. At the beginning, participants were told that they were going to listen to 15 English sentences one after another and answer either "true", "false", or "I don't know". "I don't know" was given when a sentence was not heard clearly, or when it was heard and understood but participants were not sure whether it was true or false. They were also told that the session was not for testing their English competence but for finding out how well Malaysian and Swedish speakers of English understand each other. After checking the sound volume of the headset with a sample sentence, participants listened to each sentence only once, played by the Praat software, and gave their answers at their own pace. Most of the time, they said their answers within the 7-second long interval between sentences, but some participants hesitated longer than that for some sentences. In such cases, we paused so participants could have enough time to answer. Participants listened to the sentences through the headset, and the sound also came out from a small speaker attached to the laptop computer for audio-recording both sentence reading by the speaker participants and the listeners' answers for analysis.

After the listening test was over, we invited participants to review the sentences that they listened to by re-listening and reading them on a paper, and circle words that were difficult to hear. When a participant circled a word, we asked why it was difficult. This time we also sorted the "I don't know" answer into two categories by taking participants' own clarifications: one for those not understanding pronunciations and the other for those not knowing whether the sentences were true or false. We also audio-recorded this second session for analysis.

Data analysis procedure

We recoded true/false/I don't know answers for the 15 statements from both Malaysian and Swedish listeners into two categories – "understood" and "not understood". We counted the frequencies of the first category – "understood" for all the statements as well as consonant cluster/long vowel/stress pattern statements and converted into percentages.

In addition, we measured reaction times (RT) or response latencies for the understood answers. The beginning point of the measure was at the end of the sentence reading, the ending point was the moment when the listener's answer began, and the measuring unit was seconds (s). We calculated the mean values of RTs for the 15 statements and consonant cluster/long vowel/stress pattern sentences.

To triangulate and complement the results from the true/false test, we analyzed participants' own accounts and the words they indicated as difficult to understand. We coded them into the three phonetic features that the speakers may have failed to realize clearly from the listeners' perspectives, and therefore can be considered to indicate the degrees of the intelligibility of Malay English and Swedish English pronunciations.

Moreover, instead of transforming the data to meet with the criterion of normality for inferential tests, we performed robust statistical tests with the R WRS2 Package (Wilcox & Schönbrodt, 2015), which use 20% trimmed means and do not require normal data distribution.

FINDINGS AND DISCUSSION

Which group struggled more? Do the contrasting phonetic features of Malay English and Swedish English affect listeners?

Overall, the mean percentage of Malaysian listeners' answers to all the fifteen sentences that they gave with understanding was 82.8% (SD=10.7), while that of Swedish listeners was 53.78 % (SD=17.15) (see Figure 1).

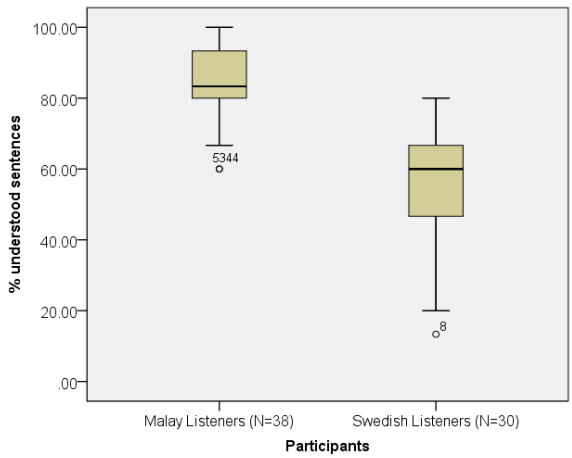

Figure 1. Percentages of the statements understood by Swedish and Malaysian listeners

A robust independent-samples *t*-test shows that the difference between Malaysian listeners and Swedish listeners was significant: the trimmed mean difference = 28.98, the 95% CI = 22.64, 35.32, t (24) = 8.15, p = 0, d = 2.03. We then analyzed the listeners' performances in terms of the three phonetic variables—stress patterns, consonant clusters, and vowel length. Firstly, 90% (SD=15.94) of Malaysian listeners' answers were given with understanding, for the sentences intended to check how clearly stress patterns were realized in the speakers' pronunciation. For the same sentences, only 57.33% (SD=25.04) of Swedish listeners' answers were made with understanding (see Figure 2). The difference between the two listener groups shown by a robust independent-samples *t*-test was significant: the trimmed mean difference =37.22, the 95% CI = 23.54, 48.90, t (25) = 6.56, p = 0, d = 1.557.

Secondly, to the sentences intended to check the intelligibility of consonant clusters pronunciations of the speakers, Malaysian listeners gave 76.36% (SD = 15.32) of their answers with understanding, while Swedish listeners gave only 32.67 % (SD = 19.29) of their answers with understanding (see Figure 3). A robust independent-samples *t*-test indicated that the difference between these percentages was significant: the trimmed mean difference =44.44, the 95% CI = 36.5, 52.39, t (34) = 11.37, p = 0, d = 2.508.

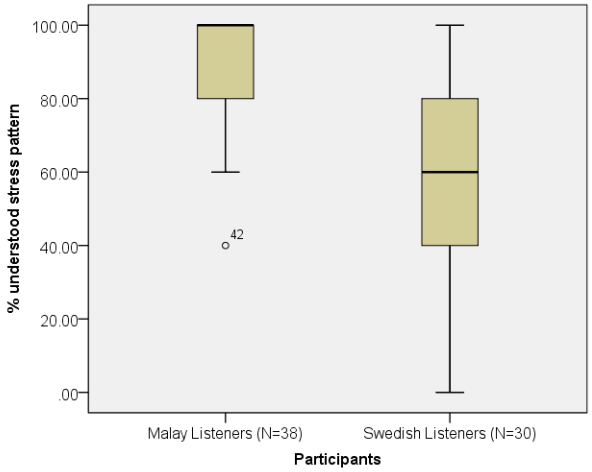

Figure 2. Percentages of the stress pattern statements understood by Swedish and Malaysian listeners

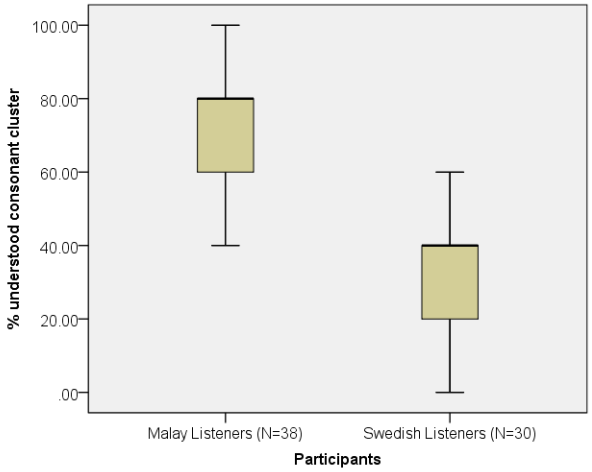

Figure 3. Percentages of consonant cluster statements understood by Swedish and Malaysian listeners.

Thirdly, 82.11% (SD=16.63) of Malaysian listeners' answers and 71.33% (SD=17.95) of Swedish listeners' answers were made with understanding for the sentences we created to test how intelligible the speakers' pronunciations were in terms of vowel length realization (see Figure 4). A robust independent-samples *t*-test indicates that difference between the two listener groups was, again, statistically significant, but with a moderate effect size: the trimmed mean difference =11.94, the 95% CI = 3.92, 19.96, t (35) = 3.03, p = 0.005, d = 0.623.

The measurement of reaction times (RT) by the two listener groups, as seen in Figure 5, also conformed to the comparison of the scores of the understood statements. As mentioned in the method section, we measured RTs only for the cases where listeners understood sentences. We firstly calculated the mean values of RTs for all the 15 sentences

taken by Malaysian and Swedish listeners. On average, the RTs of Malaysian listeners and Swedish listeners were 1.12 s (SD=0.94) and 1.6 s (SD=1.69), respectively, and the outcome of a robust independent-samples *t*-test shows that difference between the two groups was significant: the trimmed mean difference = -0.25, the 95% CI = -0.38, -0.12, t(226) = 3.75, $p = 0.0002$, $d = 0.365$.

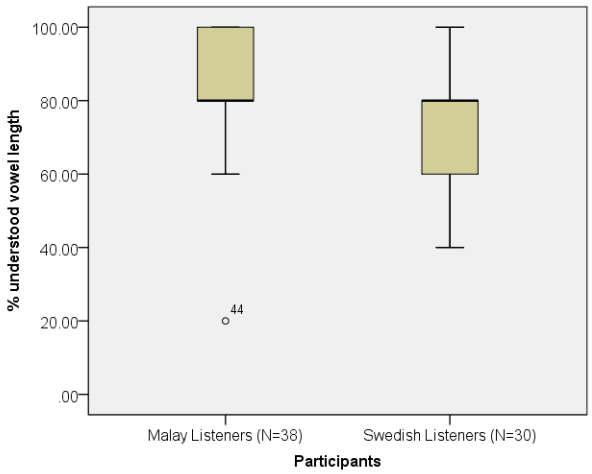

Figure 4. Percentages of the vowel length statements understood by Swedish and Malaysian listeners

However, not all the RTs for the three phonetic features clearly reveal the tendency found in the scores of understood statements. Although Malaysian listeners needed less processing times than Swedish listeners, independent-samples *t*-tests revealed that only the difference for consonant cluster sentences was statistically significant: the trimmed mean difference = -0.75, the 95% CI = -1.22, -0.27, t(37) = 3.16, $p = 0.003$, $d = 0.631$, while the other two cases were not: the trimmed mean difference = -0.18, the 95% CI = -0.44, 0.07, t(69) = 1.45, $p = 0.15$, for stress pattern, and the trimmed mean difference = -0.14, the 95% CI = -0.29, 0.005, t(135) = 1.91, $p = 0.06$, for vowel length sentences.

Nevertheless, the results overall show that Malaysian listeners understood Swedish-accented English quite well, while Swedish listeners struggled considerably to understand Malay-accented English. For Malaysian listeners, the way of the Swedish speaker' realizing word stress pattern, consonant cluster, and vowel length did not appear to affect their understanding of the speaker seriously. For Swedish listeners, on the other hand, all of the three phonetic features in Malay-accented English seemed to be difficulty factors. In particular, the consonant cluster pronunciation of the Malaysian speaker seemed to be most problematic for Swedish listeners. We look further at these possibilities in the rest of this result section by making within-group comparisons and analyzing the listener participants' own accounts.

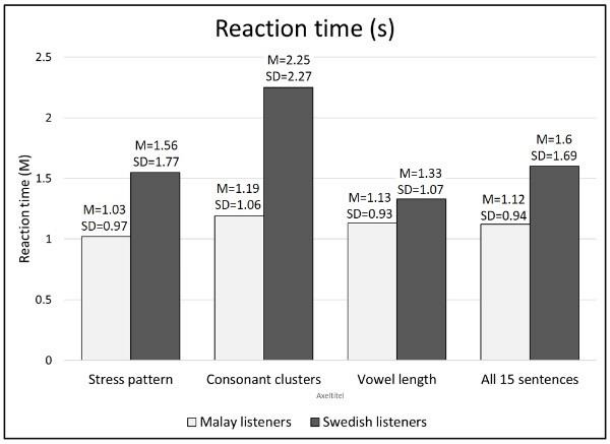

Figure 5. Reaction times for sentences understood by the two listener groups

Further analysis and discussion on difficulty factors

Firstly, as shown in Figure 6, we made within-group comparisons for stress pattern, consonant cluster, and vowel length statements. Within Malaysian listeners, understanding the consonant cluster percentage was lowest, vowel length next, and stress pattern highest. A robust repeated-measures

ANOVA test and *post hoc* tests (see Table 1) indicate that differences between the three phonetic features were all significant: F(2,46) = 18.74, p = 0, η_p^2 = .27. Therefore, although Malaysian listeners understood the Swedish speaker quite well, they felt some difficulty in understanding his consonant cluster pronunciation, significantly more than with the other two features.

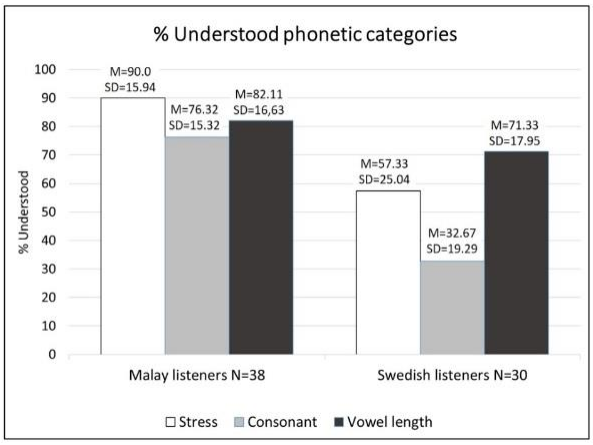

Figure 6. Percentages of understood stress pattern, consonant cluster and vowel length statements within Malaysian and Swedish listeners

Table 1. *Post hoc* analysis of difficulty factors among Malaysian listeners (N=38) as presented in understood answers

	φ	CI lower	CI upper	p	α
stress pattern vs. consonant cluster	15	8.29	21.71	.000	.017
stress pattern vs. vowel length	8.33	1.29	15.37	.006	.025
consonant cluster vs. vowel length	-5.83	-12.66	0.99	.038	.050

Meanwhile, within Swedish listeners, the percentage of understood consonant cluster sentences was again the lowest. There was a reverse in the ranking between stress pattern and vowel length when compared with the Malaysian listener within-group data. A robust repeated-measures ANOVA test and post hoc tests (see **Table 2**) show that the differences between the three phonetic features in the ranking were all significant: F(2.28) = 44.91, p = 0, η_p^2 = .849.

Table 2. *Post hoc* analysis of difficulty factors among Swedish listeners (N=30) as presented in understood answers

	φ	CI lower	CI upper	p	α
stress pattern vs. consonant cluster	24.44	10.54	38.35	.000	.025
stress pattern vs. vowel length	-13.33	-26.34	-0.32	.015	.050
consonant cluster vs. vowel length	-37.78	-45.05	-30.51	.000	.017

In addition, we also made within-group comparisons for words that Malaysian and Swedish listeners indicated as difficult by circling them. Table 3 presents the words circled by listener participants and Figure 7 shows the mean numbers of these words. Based on what we intended with these words as well as participants' own accounts, we categorized them into stress pattern, consonant cluster, and vowel length factors. For example, we used the word "bird" to see if vowel length can be an intelligibility-hurting factor, and participants' remarks agreed with our intention. Occasionally, participants pointed out issues different from what we originally intended with certain words. For example, we thought the word "instrument" would involve a stress pattern issue, but some Swedish participants said that they could not understand the Malaysian speaker because she missed the last consonant. In such cases, we took participants' own reasoning more faithfully for categorization.

We took two sets of a robust repeated measures ANOVA test and three *post hoc* tests for the words accounted as difficult by the two groups of listeners. Within Malaysian listeners, F(2,41) = 9.31, p = .00068, η_p^2 = .208, and the three *post hoc* tests in Table 4 confirmed that consonant cluster and vowel length in Swedish English gave some difficulties for Malaysian listeners, while stress pattern was not problematic. On the other hand, within Swedish listeners, F (2, 34) = 17.21, p = .00001, η_p^2 = .567. The *post hoc* results in Table 5 confirm that consonant cluster in Malay-accented English was most difficult for Swedish listeners. Stress pattern was more difficult than vowel length, but it was not statistically significant.

Table 3. Words indicated as difficult by the two listener groups

Categories	Circled by Malaysian listeners	Circled by Swedish listeners
Word stress pattern	military, semester	animal, military, only, period, salad, sausage, vegetarians
Consonant cluster	boxers, ducks, fists, glasses, smallest	boxers, ducks, eggs, elephant, fact, fast, fists, instrument, kids, lakes, most, nest, ponds, smallest, strike, textbooks, trumpet
Vowel length	birds, birth, leaves, often, peace, read, school, seats, see	birds, birth, feel, feet, floor, peace, read, school, seats, steel

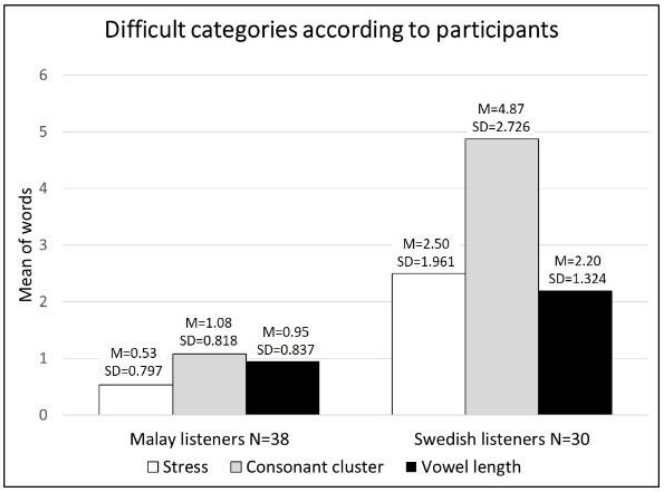

Figure 7. Mean numbers of the stress pattern, consonant cluster, and vowel length words indicated as difficult by Malaysian and Swedish listeners

Table 4. *Post hoc* comparisons for words indicated by Malaysian listeners (N=38) as difficult

	φ	CI lower	CI upper	*p*	α
stress pattern vs. consonant cluster	0.58	-0.94	-0.23	.000	.017
stress pattern vs. vowel length	-0.47	-0.77	-0.06	.006	.025
consonant cluster vs. vowel length	0.13	-0.19	0.44	.310	.050

Table 5. *Post hoc* comparisons for words indicated by Swedish listeners (N=30) as difficult

	φ	CI lower	CI upper	*p*	α
stress pattern vs. consonant cluster	-2.16	-3.43	-0.91	.000	.025
stress pattern vs. vowel length	0.11	-0.86	1.08	.765	.050
consonant cluster vs. vowel length	2.56	1.57	3.54	.000	.017

Furthermore, we had a closer look at listener participants' own accounts of their struggles. Regarding consonant cluster (the most problematic phonetic feature for both listener groups), some level of difficulty that Malaysian listeners felt with consonant clusters in Swedish English seemed to be that realizing every sound in them was different from how Malaysian English speakers pronounce them (skipping sounds). For example, a Malaysian listener, pointing at the word *fists*, told us that the Swedish speaker's pronunciation sounded strange. On the other hand, it clearly emerged among Swedish listeners that their listening comprehension dropped off when some sounds in consonant clusters were missing. For example, some comments were: "The woman who spoke sort of swallowed the words…Often, like this one, *textbooks*, or (Pointing at *fast*) it was difficult to hear because *t* was missing…yeah, there are a lot of t's missing".

As for stress pattern and vowel length words, insufficient lengthening of long (tense) vowels and inconsistent stress pattern were mentioned by Swedish listeners as what negatively affected their understanding of the Malaysian speaker to a large extent. The overall response from Swedish listeners is aligned with our own assessment of the Malaysian speaker. That is, we observed that she skipped many word-final consonants and hardly pronounced long/tense vowels with enough length to discern them from short/lax vowels. In addition, to our perception, her word stress in polysyllabic words was either unclear, displaying little difference between stressed and unstressed syllables, or evidently on the wrong syllable. For example, she stressed the second syllable of the word "only" as [onˈliː], and the first syllable of the word "vegetarians" as [ˈvɛʤɪˌtərɪəns].

The findings reveal that, for Malaysian English users who do not depend on the three phonetic features tested in the study, clear realization of these features had little negative impact on their listening comprehension. Meanwhile, for Swedish English users who heavily use the same features as perceptual cues, a foreign accent that does not realize them clearly was greatly challenging. These findings support previous studies that suggest stress patterns, consonant clusters, and vowel quantity (tense-lax) contrast as crucial features for intelligible speech.

In the dimension of word stress pattern, our result is coherent with Field (2005), Hahn (2004), Atsuko et al. (2008), Rajadurai (2006), Tiffen (1992), and Zielinski (2008). It challenges Jenkins' (2002) conclusion that word stress placement is not a phonetic core feature in the context of NNS-NNS communication. As seen in the result, the tendency of Malay-accented English—stressing the last syllable of the word (Yong 2001) or otherwise changing word stress patterns—seems to be a difficulty factor for Swedish listeners who depend on stress pattern to recognize words. We compared only two NNS groups, and a problematic feature in Malay-accented English to Swedish listeners may not be an issue at all to listener groups from other L1 backgrounds. We therefore cannot claim that word stress should be regarded as a feature of EIL crucial for intelligibility in general. Nevertheless, based on the findings, it can be suggested that word stress pattern (i.e., which syllable should carry primary stress in a word) needs to be considered in developing the description of intelligible EIL pronunciation.

In addition, the findings of the study confirm studies that showed that omitting sounds in consonant clusters decreases intelligibility (Jenkins, 2002; Lesley, 2014; Sewell, 2016; Suenobu et al., 1992; Tajima et al., 1996; Zhang, 2015). In fact, consonant cluster emerged as the most influential factor for intelligibility among the three tested features. Some Malaysian participants expressed difficulty for understanding clearly realized consonant clusters. Overall, however, our cross-linguistic comparison reveals that for L2 users who tend to simplify consonant clusters, hearing all sounds is not so difficult, but for L2 users who do not skip sounds, consonant reduction was a great difficulty factor. The result makes us cautious about Sewell's (2016) argument that consonant reduction is rather universal among different varieties of English and there are problematic and non-problematic cluster simplifications. That is, as reviewed previously he suggested that the consonant deletion in mono- and disyllabic words, or omitting consonants carrying grammatical information (e.g., omitting the last [-d] from "*learned*") is problematic, while skipping the final consonants in multi-syllable words (e.g., omitting the last [-t] from *department*) is less detrimental since the missing segment constitutes a minor part of the whole word. However, Swedish listeners in our study struggled not only with missing consonant in mono or disyllabic consonant words, but also when the final consonants in multi-syllabic words (e.g. [-t] in "*elephant*" or "*instrument*") were not heard.

Moreover, our data, which show that the insufficient duration of long/tense vowels can be a negative factor for intelligibility, are aligned with a body of research that revealed that the contrast between long and short vowels should be maintained for intelligible EIL pronunciation (Bent et al., 2008; Jenkins, 2002; Quené & van Delft, 2010; Rajadurai, 2006; Smith et al., 2003; Tajima et al., 1996). Swedish listeners were confused when long vowels of the Malaysian speaker such as /ɜː/ in "bird" or /iː/ in "seats" were shorter than expected. As reviewed previously, whether or not Malay English realizes contrast in vowel quantity can be debatable (see Baskaran, 2008; Pillai et al., 2010; Yong, 2001), and the result of our study does not provide a supporting point for either side. However, it shows that, whether or not Malay English maintains the vowel quantity (tense-lax) contrast, the duration of its long vowel sounds was not long enough for its intelligibility, at least against the perception of Swedish listeners.

CONCLUSION

In the study, we compared the intelligibility of Malay-accented English and Swedish-accented English against the mutual perceptions of Malaysian and Swedish listeners. The result shows that Malaysian listeners could more easily understand the 15 English statements spoken by the Swedish speaker, compared to the Swedish listeners listening to the Malaysian speaker. Both groups of speakers have their own accents which are different from any of NS varieties. However, when evaluating them against the intelligibility principle, the contrast between them seems to be related to the degree of realization of phonetic features affecting the intelligibility of L2 speech (Derwing & Munro, 2015).

After this general conclusion, we want to point out some possible biases of our study. Firstly, we looked at the intelligibility of Malay English and Swedish English only in the aspect of pronunciation, and other possible parameters, such as pragmatic strategies or grammaticality, were not concerned. Secondly, we compared two groups of EIL users, one of which had an L1 that is genetically closely related to English (Germanic origin) with a phonologic structure that is quite similar to native varieties of English. Accordingly, a reason for the asymmetric result of the study could be that Malaysians in general may have been exposed substantially to Western varieties of English (e.g.,

American or British varieties) while the Swedish may not be familiar with Asian English varieties. Thirdly, we began from hypothesizing three phonetic variables as the intelligibility-affecting factors, but it should be noted that other phonetic features of Malay English and Swedish English can affect their intelligibility. Fourthly, although four raters evaluated the speaker participants as having typical Malay English and Swedish English accents (see the method section), it can be questioned how well one speaker from each linguistic group represents the pronunciations of the whole group.

Regarding the second point of the limitations of our study, we wish to suggest an implication for teaching pronunciation to L2 learners based on the findings. As Jenkins (2002), Derwing and Munro (2015) and many other researchers noted, the concept of intelligibility involves the mutual responsibility of both speakers and listeners for successful communication. L2 pronunciation teaching should centrally involve raising awareness of both speakers' and listener's responsibility. Learners should be facilitated to acquire phonetic features found to be important for international communication, such as those examined in the present study. Meanwhile, they also need to be exposed to different English accents, particularly to those that considerably differ from their own, to develop perceptual competence for a wide range of EIL pronunciations.

Finally, we suggest future studies examine the intelligibility of EIL pronunciation against the perceptions of people representing diverse varieties of English.

ACKNOWLEDGEMENTS

We thank Mr. Robert Boxwell for providing language feedback and our participants for generously sharing their time.

REFERENCES

Abelin, Å., & Thorén, B. (2015). What affects recognition most – wrong word stress or wrong word accent? Paper presented at the Fonetik, Lund.

Abelin, Å., & Thorén, B. (2016). Identification of stress, quantity and tonal word accent in Swedish. Paper presented at the New Sounds Conference Aarhus, Denmark.

Baskaran, L. 2008. Malaysian English: Phonology. In Mesthrie, R. (ed.), *Varieties of English: Africa, South and Southeast Asia* (278-291). Berlin: Mouton de Gruyter.

Bent, T., Bradlow, A. R., & Smith, B. L. (2008). Production and perception of temporal patterns in native and non-native speech. *Phonetica, 65*(3), 131-147. doi:10.1159/000144077

Best, C. T. (2016). Perceptual assimilation reflects native phonetic-phonological mapping: perceiving and adopting to regional accent differences in native English. Paper presented at the New Sounds, Aarhus, Denmark.

Boersma, P., & Weenink, D. (2014). Praat: Doing phonetics by computer. Version 5.4.04. Retrieved from http://www.praat.org/

Brown, A. (1988). The staccato effect in the pronunciation of English in Malaysia and Singapore. In J. Foley (Ed.), *New Englishes: The case of Singapore* (pp. 115-128). Singapore: Singapore University Press.

Davidsen-Nielsen, N., & Harder, P. (2001). Speakers of Scandinavian languages: Danish, Norwegian, Swedish. In M. Swan & B. Smith (Eds.), *Learner English: A teacher's guide to interference and other problems* (2nd ed., pp. 21-36). Cambridge: Cambridge University Press.

Derwing, T. M., & Munro, N. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. Amsterdam: John Benjamins.

Elert, C.-C. (1964). *Phonologic studies of Swedish quantity*. Uppsala, Sweden: Almqvist & Wiksell.

Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly, 39*(3), 399-423. doi:10.2307/3588487

Garlén, C. (1988). *Svenskans fonologi*. Lund, Sweden: Studentlitteratur.

Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly, 38*(2), 201-223. doi:10.2307/3588378

Jenkins, J. (2002). A sociolinguistically based, empirically researched pronunciation syllabus for english as an international language. *Applied Linguistics, 23*(1), 83-103. doi:10.1093/applin/23.1.83

Kashiwagi, A., & Snyder, M. (2008). American and Japanese listener assessment of Japanese EFL speech: Pronunciation features affecting intelligibility. *The Journal of Asia TEFL, 5*(4), 22-47.

Kashiwagi, A., Snyder, M., & Craig, J. (2006). Suprasegmentals vs. segmentals: NNS phonological errors leading to actual miscommunication. *JACET Bulletin, 43*, 43-57.

Lesley, J. (2014). A phonological and prosodic analysis of English pronunciation by Japanese learners. *Hawaii Pacific University TESOL Working Paper Series, 12*, 18-31.

Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly, 39*(3), 369-377. doi:10.2307/3588485

McKay, S. L. (2009). English as an international language: Where we are and where we need to go. *Journal of English as an International Language, 5*, 27-54.

Munro, M., & Derwing, T. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech, 38*(3), 289-306. doi:10.1177/002383099503800305

Munro, M., & Derwing, T. (2013). The foundations of accent and intelligibility in pronunciation research. *Language Teaching, 44*(3), 316-317. doi:10.1017/S0261444811000103

Phoon, H. S., Abdullah, A. C., & Maclagan, M. (2013). The Consonant realizations of Malay-, Chinese- and Indian-influenced Malaysian English. *Australian Journal of Linguistics, 33*(1), 3-30. doi:10.1080/07268602.2013.787902

Pillai, S., Mohd. Don, Z., Knowles, G., & Tang, J. (2010). Malaysian English: An instrumental analysis of vowel contrasts. *World Englishes, 29*(2), 159-172. doi:10.1111/j.1467-971X.2010.01636.x

Quené, H., & van Delft, L. E. (2010). Non-native durational patterns decrease speech intelligibility. *Speech Communication, 52*(11–12), 911-918. doi:10.1016/j.specom.2010.03.005

Rajadurai, J. (2006). Pronunciation issues in non-native contexts: A Malaysian case study. *Malaysian Journal of ELT Research, 2*, 42-59.

Sewell, A. (2016). *English pronunciation models in a globalized world: Accent, acceptability and Hong Kong English*. Oxen: Routledge.

Smith, B., Bradlow, A. R., Bent, T., Solé, M. J., Recasens, D., & Romero, J. (2003). Production and perception of termporal contrasts in foreign-accented English. Paper presented at the XVth International Congress of Phonetic Science.

Suenobu, J., Kanzaki, K., & Yamane, S. (1992). An experimental study of intelligibility of Japanese English. *International Review of Applied Linguistics in Language Teaching, 30*, 146-156.

Tajima, K., Port, R., & Dalby, J. (1996). Effects of temporal correction on intelligibility of foreign-accented English. *Journal of Phonetics, 25*(1), 1-24. doi:10.1006/jpho.1996.0031

Thomson, R. I. (2014). Myth 6: Accent reduction and pronunciation instruction are the same thing. In L. Grant (Ed.), *Pronunciation myths: Applying second language research to classroom teaching* (pp. 160-187). Michigan: University of Michigan Press.

Thorén, B. (2003). Can V/C-ratio alone be sufficient for discrimination of V:C/VC: in Swedish? A perception test with manipulated durations. Paper presented at the Fonetik 2003, Umeå, Sweden.

Tiffen, B. (1992). A study of the intellgibility of Nigerian English. In A. van Essen & E. I. Burkard (Eds.), *In homage to W. R. Lee: Essays in English as a foreign or second language* (pp. 255-259). Berlin, Germany: Foris.

Wilcox, R. R., & Schönbrodt, F. D. (2015). The WRS package for robust statistics in R (Version 0.27.5). Retrieved from https://github.com/nicebread/WRS

Yong, J. Y. (2001). Malay/Indonesian speakers. In M. S. a. B. Smith (Ed.), *Learner English: A teacher's guide to interference and other problems* (2nd ed., pp. 279-295). Cambridge, England: Cambridge University Press.

Zhang, L. (2015). An empirical study on the intelligibility of English Spoken by Chinese university students. *Chinese Journal of Applied Linguistics, 38*(1), 36. doi:10.1515/cjal-2015-0002

Zielinski, B. W. (2008). The listener: No longer the silent partner in reduced intelligibility. *System, 36*(1), 69-84. doi:10.1016/j.system.2007.11.004