Effect of altering three phonetic features on intelligibility of English as a lingua franca: a Malaysian speaker and Swedish listeners

Hyeseung Jeong, Bosse Thorén & Juliana Othman

To cite this article: Hyeseung Jeong, Bosse Thorén & Juliana Othman (2018): Effect of altering three phonetic features on intelligibility of English as a lingua franca: a Malaysian speaker and Swedish listeners, Asian Englishes, DOI: 10.1080/13488678.2018.1536817

To link to this article: https://doi.org/10.1080/13488678.2018.1536817

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

Published online: 01 Nov 2018.

Submit your article to this journal

Article views: 189

View Crossmark data
Effect of altering three phonetic features on intelligibility of English as a lingua franca: a Malaysian speaker and Swedish listeners

Hyeseung Jeong, Bosse Thorén and Juliana Othman

ABSTRACT
Our previous study examined the mutual intelligibility of Malaysian English to Swedish listeners and Swedish English to Malaysian listeners. The results showed that Swedish listeners did not understand the Malaysian speaker well. In the present study, the Malaysian speaker was trained to alter her realization of the word stress, consonant clusters and long vowels in a way that previous research has found intelligible for both native and non-native English speakers. The audible and measurable alteration significantly increased the intelligibility of the speaker for Swedish listeners. This indicates that the three phonetic features are important for intelligibility in international contexts and suggests including the word stress in the Lingua Franca Phonetic Core. Moreover, we discuss that Malaysian English being a dialect and Swedish English being a similect may be relevant to their mutual intelligibility and relate the discussion to teaching English pronunciation in countries where English has been localized.

ARTICLE HISTORY
Received 27 April 2018
Accepted 12 October 2018

KEYWORDS
English as a lingua franca; intelligibility; intelligible pronunciation; Malaysian English; Swedish listeners; Lingua Franca Phonetic Core

Introduction
In using English as a lingua franca (ELF), the intelligibility principle has emerged as an alternative to the native speaker norm (Jenkins, 2000, 2015; Levis, 2005). According to the principle, a clear realization of a limited number of certain phonetic or phonological features suffices for being understood.

As a way to test the intelligibility principle, our previous study (Jeong, Thorén, & Othman, 2017) looked at mutual intelligibility between Malaysian-accented English and Swedish-accented English. Specifically, we wanted to test the effects of the three phonetic features – word stress, consonant cluster and vowel length – on the mutual intelligibility; the features that literature suggests are important for intelligible pronunciation, although in mixed and inconclusive ways (e.g., J. Field, 2005; Hahn, 2004; Jenkins, 2002). The results showed that Swedish English was more intelligible to Malaysian listeners than Malaysian English to Swedish listeners. Both statistical and interview data indicated that an inconsistent stress pattern, simplification of consonant
clusters and shortened long-vowel sounds in Malaysian English seemed to make it unintelligible to Swedish listeners to a great degree. On the contrary, the rather clear realization of the three phonetic features in Swedish English did not appear to hurt its intelligibility against Malaysian listeners’ perception, although it is not aligned with Malaysian English phonetics in general. Our tentative conclusion, therefore, was that a clear realization of the English stress pattern, consonant clusters and long vowels may be beneficial for, or at least does not hurt, intercultural communication. When we mention ‘clear realizations’ for the stress pattern, we refer to the Inner Circle or Standard English rules, which are sometimes referred to for conceptualizing intelligible ELF pronunciation (see Jenkins, 2002). This, however, should be distinguished from the native speaker norm that tells us that second-language (L2) speakers should aim for every single Standard English rule.

This current study is a follow-up of our previous one (Jeong et al., 2017). It aims to find out whether the intelligibility of Malaysian English for Swedish listeners could be improved if its word stress patterns, consonant clusters and long vowels are altered for clear realization, while all the other features are maintained. The study also attempts to determine whether the alterations of the three phonetic features have differential effects and whether they lead to a more balanced mutual intelligibility of Malaysian English and Swedish English. Later, we will relate our findings to Jenkins’ (2000, 2002, 2015) syllabus of phonetic core features of ELF.

The research questions are:

(1) Do the alterations of the word stress pattern, consonant clusters and vowel length increase the intelligibility of Malaysian English to Swedish listeners’ perception?

(2) If the answer for the first question is yes, do the altered realizations of the three phonetic features contribute equally to the intelligibility of Malaysian English for Swedish listeners, or do they show differential effects?

(3) Do the alterations of the three phonetic features result in a balanced mutual intelligibility between Malaysian and Swedish English?

**Literature review**

**Phonetic core features of ELF**

Together with the overarching argument that L2 speakers need to aim for an intelligible pronunciation rather than native-likeness, the intelligibility principle denotes that a speech can be either more or less intelligible or unintelligible. Jenkins’ (2000, 2002) syllabus of the Lingua Franca Core (LFC) is a comprehensive specification of whether certain English phonetic features are important for making a speaker intelligible and therefore can be a guideline for assessing intelligibility in using ELF. The syllabus includes several details as follows. First, among the variations of a phoneme, there can be one that is more easily understood than the other. For example, between the two allophones of /t/ in ‘water’, [t] is more likely to be intelligible to L2 speakers than flapped [ɾ] found in American English. Second, some substitutions of certain phonetic features, such as /θ/ and /ð/, can be tolerated and do not cause considerable
misunderstandings. Most crucially, Jenkins lists segmental and prosodic features of English that one needs to pronounce clearly in order for the international interlocutor to understand well. For example, consonant clusters in general need to be realized clearly. In particular, any sound in the initial consonant cluster (e.g. in STReet or DReam) cannot be deleted, and omitting some sounds in middle and final consonant clusters is only possible in accordance with the syllable structure rules of first-language (L1) varieties (e.g. not ‘fatsheet’ or ‘facteet’ but ‘facsheet’ for pronouncing ‘faCTSHeet’).

Jenkins’ (2000, 2002) description of the LFC received criticism, most of which is discussed in Derwing and Munro (2015). First, it is not clearly evidence based. It was based on a small amount of data, and the criteria for classifying core and non-core features were unclear and have not been attested through further research. In fact, other researchers have questioned some patterns included in her specification, such as regarding epenthesis as permissible, or, as discussed further in the following section, suggesting that the word stress pattern is not a core feature. One more critical thing that impairs the credibility of Jenkins’ work is the exclusion of native interlocutors in her empirical data, as if ‘lingua franca speakers will not talk to native speakers’ (Derwing & Munro, 2015, p. 144).

Despite issues in it, Jenkins’ LFC syllabus itself can be undoubtedly useful for operationalizing the intelligibility principle in teaching and assessing pronunciation. For example, through their experiment conducted in Iran, Rahimi and Ruzrokh (2016) found that a pronunciation instruction using Jenkins’ syllabus had a very positive effect on English learners’ pronunciation intelligibility as well as on their recognition intelligibility (i.e. comprehension). Findings like these suggest that the LFC syllabus can provide descriptive guidelines (as opposed to a prescriptive syllabus), based on substantial empirical data, and should be continuously tested in diverse contexts.

**Stress pattern, consonant clusters and vowel length: core or non-core features?**

In Jenkins’ (2000, 2002, 2015) LFC syllabus, clear realization of consonant clusters and contrasting vowel length are counted as core features, while word stress is classified as a non-core feature.

First, Jenkins (2002) listed as a core of the ELF phonology not omitting sounds in consonant clusters except for those permitted in L1 English phonologies, and this already has been confirmed by many empirical studies (e.g., Sewell, 2015; Suenobu, Kanzaki, & Yamane, 1992; Tajima, Port, & Dalby, 1996). For example, Tajima et al. (1996) manipulated a Chinese speaker’s short phrases by inserting missing consonants in the consonant clusters and compared this to native speaker listeners’ perception of the original and manipulated speeches. They found that the intelligibility of the manipulated speech was significantly higher than the original version. Suenobu et al. (1992) and Zhang (2015) tested the intelligibility of Japanese and Chinese speakers’ pronunciations against native and international listeners, respectively. The results of both studies showed that simplifying consonant clusters was one of the factors hampering intelligibility. In addition, Sewell (2015) reports that omitting or changing the quality of sounds in consonant clusters can cause communication breakdown, even for intranational interaction in Hong Kong.
Secondly, the LFC syllabus also suggests that pronouncing short and long vowel sounds distinctively is a core feature of ELF. Jenkins (2000) notes that the quality of vowel phonemes can differ, but their quantity is consistent and stable across different varieties of English. Some empirical studies, such as Bent, Bradlow, and Smith (2008), Quené and van Delft (2010), Smith et al. (2003) and Tajima et al. (1996), have observed the effect of contrasting long and short vowels on intelligibility. On the contrary, Derwing and Munro (2015) have presented a different view. They acknowledge that not differentiating minimal pairs, like beet/bit, requires a lot of effort for the listener to understand and is therefore a factor that negatively affects intelligibility. However, they argue that the fundamental difference between vowels in the minimal pair is qualitative; mainly, while one is tense, the other is lax, signalled by spectral differences rather than temporal. Therefore, more empirical data would be necessary to confirm the inclusion of contrasting vowel length in the syllabus.

As for stress patterns, some studies found that stress placement is an important factor for both native and non-native speaker listeners’ comprehension (J. Field, 2005; Kashiwagi, Snyder, & Craig, 2006; Saito & Shintani, 2016). In J. Field (2005), non-native speaker listeners from six different L1 backgrounds (Korean, Japanese, Mandarin, Spanish, Portuguese and Italian) and English native speaker listeners reacted similarly to Japanese speakers’ misplaced lexical stress. Likewise, Kashiwagi and Snyder (2008) revealed that Japanese listeners as well as American listeners (English native speakers) struggled to comprehend Japanese learners of English when they misplaced word stress. Their findings, in fact, made them reject their own previous study (Kashiwagi et al., 2006), in which they argued that native speakers seemed to use the stress pattern as an important perceptual cue, but non-native speakers did not. In addition, Saito and Shintani (2016) present results indicating that both North Americans’ and Singaporeans’ perceptions of the comprehensibility of the L2 speech were influenced by whether or not it locates word stress clearly. However, Jenkins (2002, 2015) excluded the stress pattern, both at word and sentence levels, from the core. Jenkins states that the stress pattern is critical at a word level and at least important at a sentence level for the native speaker/listener, but not for the ELF speaker/listener. Recently, Jenkins acknowledged that native speakers are also participants of ELF communications, but the LFC still regards the word stress pattern as non-core (Jenkins, 2015). In light of accumulated findings, word stress appears to be a factor that influences intelligibility, with which further research can make a strong argument for including the stress pattern in the core features in the LFC.

Stress pattern, consonant cluster and vowel length in Malaysian English

A dialect of English is, either as a native or non-native variety, a localized form of the language that arises from ‘an English-using community, sufficiently large and sufficiently stable as a community’ (Kachru, 1992, p. 34). Malaysia is a multi-ethnic, multi-lingual country, and there are different varieties of English accents among Malaysians (e.g., Baskaran 2004, 2008). Nevertheless, these intranational varieties share certain generic phonetic features (Baskaran 2004, 2008; Mesthrie & Bhatt, 2008; Yong, 2001), which, together with other linguistic features, endow Malaysian English the status of a dialect of World Englishes. The following paragraphs review the description of stress
pattern, consonant cluster and vowel length of Malaysian English, and include our own observations.

From the descriptions of phonetic features of Malaysian English, we summarize that Malaysian English usually puts stress on the ultimate syllable of words, which is assumed to be a trace of Malay language (e.g. carefullyLY instead of CArefully in ‘I tried to read the letter carefully’). In addition, Brown (1988), as well as Baskaran (2004, 2008), Mesthrie and Bhatt (2008) and Yong (2001), notes that unstressed syllables of Malaysian English are not reduced in duration and vowel quality to the extent that those of Inner Circle Englishes are, giving an impression of less contrast between stressed and unstressed syllables. Taken together, this makes the typical Malaysian English stress pattern partly less distinct and partly changed, compared to Inner Circle Englishes. The reduced contrasts between stressed and unstressed syllables, caused by increased equality in both length and articulation, put Malaysian English more on the syllable-timed side of the continuum of ‘stress-timed’ and ‘syllable-timed’.

Concerning consonant clusters, our own impression is that Malaysian English often has reduced consonant clusters and sometimes omitted singleton word-final obstruents. Baskaran (2004, 2008), Phoon, Abdullah, and Maclagan (2013) and Mesthrie and Bhatt (2008) give examples of three-consonant clusters reduced to two-consonant clusters as well as two-consonant clusters reduced to one (e.g. ‘as’ for ‘ask’, ‘jum’ for ‘jump’, ‘haan’ for ‘hand’). Reduction of consonant clusters takes place mostly in word-final and word-medial positions.

Malaysian English tends not to prolong vowels or syllables. Baskaran (2004) describes that this general tendency to make all vowels short is an influence of Bahasa (e.g. [bit] for ‘beat’ /biːt/ or [ful] for ‘fool’ /fuːl/). He also draws on some examples where some L1 English short vowels become long when followed by the apical consonants/n, l, r, s or ʃ/. However, he considers that these cases can possibly be lexical rather than phonological. In addition, Brown (1988) notes that Malaysian and Singaporean English replace syllable-final /p, t, k/ with a glottal stop, which in turn has a shortening effect on the preceding vowel.

Taken together, Malaysian English, with the tendencies for stress patterns and to those for vowel length, seems to have a speech rhythm that is generally syllable timed, with stress mostly on the word’s final syllable and reduced contrast between stressed and unstressed syllables. The general tendencies concerning reduction of consonant clusters are that two and three-consonant clusters are reduced, often by deleting a cluster-final obstruent.

**Method**

This study followed the same methodological procedure as the first study (Jeong et al., 2017), as described in the following.

**Speaker**

The speaker was the Malaysian female whose reading was recorded and used for the first study. She was 21 years old and was a student at an English teacher education programme. Malay is her first language and English her second. Four international
people experienced with Malaysian English speakers commented that she spoke English with a moderate Malay accent. This means that her English accent had the generic phonetic features of Malaysian English described in the literature review section, in relation to how stress pattern, consonant cluster and vowel length are realized.

**Listeners**

Twenty-one Swedish listeners were either staff or students at two Swedish universities, who did not participate in the first study. Their ages ranged from 25 to 67 years. All of the participants began to learn English at ages 2–13 years and used English on a daily or regular basis. They hardly used English to communicate with other Swedes but mostly for international communication, which is typical as speakers of an English similect (Mauranen, 2012; for further discussion on similect, refer to the Discussion and conclusion section). In academic contexts, they read materials written in English, and had written or spoken communications with international people including staff and students in and outside Sweden. Besides, they used English for participating in social media and for travelling abroad, and they had substantial English input through Television programmes. None of the participants reported hearing problems that could affect their listening performance. In the Results section, we compare this group of 21 Swedish listeners with two other groups from the first study. One of the two was a Swedish listener group (30 people) that listened to the same speaker before training (for answering research question 1), and the other was a Malaysian listener group (38 people) that listened to a Swedish speaker reading the same sentences (for answering research question 3).

**Speech material**

We used the same true/false statements created for the first study. They consisted of 5–11 words ($M = 7.07$) and were constructed in a way that they could ‘be easily determined as true or false when they are understood’ (Jeong et al., 2017, p. 45). The words intended to test the respective phonetic features are presented in bold.

**Sentences testing stress patterns**

- Vegetarians like to eat sausage salad.
- The smallest animal in Africa is the elephant.
- Military service is for women only.
- A semester is a period in schools or universities.
- A trumpet is a musical instrument.

**Sentences testing consonant clusters**

- Kids wear glasses to walk fast.
- Most birds make a nest to lay their eggs.
- Ducks often swim in lakes and ponds.
- Lots of textbooks describe facts.
- Boxers must use only their fists to strike each other.
**Sentences testing long vowels**

- We can feel with our feet when the floor is warm.
- Birds can read from birth.
- You often see leaves on trees.
- Car seats must be made of steel.
- Nobody wants peace on earth.

**Alteration of stress pattern, consonant cluster and long vowels**

The speaker altered her pronunciation of the stress pattern, consonant cluster and long vowels through a one-week training session, while maintaining other phonetic features of her English accent. The training aimed at clearly realizing all of the sounds in the consonant cluster words and sufficiently lengthening the long vowels in the long vowel words, following Jenkins’ (2002, 2015) suggestion of the intelligible pronunciations of the two phonetic features in the context of ELF. In addition, she aimed to alter her stress patterns for the stress pattern words to patterns agreeing with the Inner Circle Englishes, which previous studies found out to be important for an intelligible English pronunciation against both native and non-native listeners (e.g., J. Field, 2005).

The training was overall successful. We compared the two recorded versions according to the perception of the second author (phonetician) and to the following acoustic variables: for word stress patterns we used one temporal and one tonal measure. Duration and pitch have been found to be relevant acoustic cues for word stress in English by, for example, Fry (1958) and Eriksson and Heldner (2015). The temporal measure consisted of the vowel duration divided by the entire word duration in the syllables which should receive stress according to Inner Circle Englishes. This ratio was compared between the first and second recordings. The absolute vowel duration in the measured syllables was increased by approximately 30% on average, but the relative vowel duration was not increased significantly, since the overall speaking rate was lower in the second version.

The tonal measures were the pitch range (semitones) within the voiced part of the stressed syllable and also the pitch range divided by the syllable duration, yielding the pitch change per second or pitch slope. We must, however, bear in mind that the pitch range, pitch slope and syllable duration contribute to the impression of stress, and accordingly, with the same pitch range and increased syllable duration, the pitch slope actually decreases. However, in 12 out of 16 target words, the pitch range as well as the pitch slope in the intended stressed syllable was increased by on average 1.4 semitones and 6.7 semitones per second respectively.

For vowel length, we measured absolute durations of target vowels and also their portion of the entire word duration. This measure yielded substantially higher values for 13 target words out of 15 and a mean increase, including all target words, by 30%.

Altered consonant clusters were checked by means of auditive control as well as visual examination of spectrograms. In 8 out of 13 target words, there was an increased number of audible and visible obstruent consonants. In the word ‘textbooks’, which contains two consonant clusters, [kst] and [ks], both were improved in the second recording. On average, 0.7 obstruent consonant segments per target word were added from the first to the second recording.
Recordings and material preparation

The speaker was recorded at 48,000 Hertz in a music studio at the University of Malaya, and this was saved as a Microsoft WAVE file. By means of the software Praat (Boersma & Weenink, 2014) we then randomly organized the sentences with seven-second silence intervals between them, out of which we also created two versions of the sound file with reversed order to eliminate possible training effects.

Data collection procedure

Data collection was carried out individually in a quiet room and took 10–15 minutes. 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 answered either ‘true’, false’ or ‘I don’t know’. The target sentence and the subject’s reply were recorded to check the reaction time. After the listening test, participants reviewed the sentences written on paper, and told us which sentences they heard correctly and which ones not. We also sorted the ‘I don’t know’ answers into two categories according to participants’ own clarifications: one for those not understanding pronunciations, and the other for those that were understood but with the participant not knowing whether the sentence was actually true or false. We audio-recorded all of the sessions with individual participants.

Data analysis procedure

We compared the performance of the Swedish listeners with those from the first study comprising 38 Malaysian participants who listened to the Swedish speaker and 31 Swedish participants who listened to the same Malaysian speaker before she altered her pronunciation. We labelled the data of Malaysian and Swedish participants from the first study as ‘Malaysians’ and ‘Swedes 1’, and Swedish participants’ data from this study as ‘Swedes 2’.

For analysis, we had to exclude the sentence ‘Boxers must use only their fists to strike each other’, intended to check consonant clusters. Many participants thought its content was much more difficult to understand than the other sentences, and we came to think that their answers for it did not clearly indicate their understanding of the speaker’s pronunciation. We then sorted the answers for the remaining 14 statements and sorted them into two categories – ‘understood’ and ‘not understood’ – based on participants’ own comments given during the review session. The number of answers falling into the ‘understood’ category was counted and converted into percentages. We also measured reaction times for the answers.

As the data were not normally distributed and did not meet the assumptions for using standard parametric tests (A. Field, 2018, p. 753), we performed robust statistical tests with the R WRS2 Package (Wilcox & Schönbrodt, 2015) and non-parametric tests using SPSS for the comparisons, as advised by Larson-Hall (2015) and Turner (2014).
Results

Descriptive statistics

Here we present descriptive statistics including means and standard deviations as well as 20% trimmed means used for robust inferential statistics, before answering the three research questions.

First, Table 1 presents the means, standard deviations and trimmed means of the listeners’ answers given with understanding the 14 test sentences, and Figure 1 shows only the trimmed means. As mentioned in the Method section, ‘Malaysians’ are those

Table 1. Means, standard deviations and 20% trimmed means of listeners’ answers given with understanding.

<table>
<thead>
<tr>
<th>Sentences</th>
<th>Participants(^a)</th>
<th>Mean (%)</th>
<th>Standard deviation</th>
<th>20% trimmed mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>All sentences</td>
<td>Malaysians</td>
<td>86.47</td>
<td>10.86</td>
<td>84.04</td>
</tr>
<tr>
<td></td>
<td>Swedes 1</td>
<td>56.9</td>
<td>18.33</td>
<td>58.34</td>
</tr>
<tr>
<td></td>
<td>Swedes 2</td>
<td>86.05</td>
<td>7.31</td>
<td>85.70</td>
</tr>
<tr>
<td>Stress pattern sentences</td>
<td>Malaysians</td>
<td>90</td>
<td>15.94</td>
<td>95.26</td>
</tr>
<tr>
<td></td>
<td>Swedes 1</td>
<td>57.33</td>
<td>25.04</td>
<td>57.78</td>
</tr>
<tr>
<td></td>
<td>Swedes 2</td>
<td>96.19</td>
<td>8.05</td>
<td>100</td>
</tr>
<tr>
<td>Consonant cluster sentences</td>
<td>Malaysians</td>
<td>87.5</td>
<td>16.18</td>
<td>77.02</td>
</tr>
<tr>
<td></td>
<td>Swedes 1</td>
<td>38.33</td>
<td>23.43</td>
<td>40.28</td>
</tr>
<tr>
<td></td>
<td>Swedes 2</td>
<td>65.48</td>
<td>24.34</td>
<td>65.48</td>
</tr>
<tr>
<td>Long vowel sentences</td>
<td>Malaysians</td>
<td>82.11</td>
<td>16.63</td>
<td>83.86</td>
</tr>
<tr>
<td></td>
<td>Swedes 1</td>
<td>71.33</td>
<td>17.95</td>
<td>72.22</td>
</tr>
<tr>
<td></td>
<td>Swedes 2</td>
<td>92.38</td>
<td>11.79</td>
<td>95.56</td>
</tr>
</tbody>
</table>

Note: \(^a\)Number of participants: Malaysians = 38, Swedes 1 = 30, Swedes 2 = 21.

Figure 1. The 20% trimmed means of listeners’ answers given with understanding.
who listened to a Swedish speaker, ‘Swedes 1’ are those who listened to the Malaysian speaker before she altered her stress pattern, consonant cluster and long vowel features, and ‘Swedes 2’ are those who listened to the same speaker after she altered them. Second, Table 2 presents the means, standard deviations and trimmed means of the reaction times that the same listener participants took when giving their answers, and Figure 2 shows the trimmed means. In addition to the descriptive statistics provided here, we calculated the standard scores (z values) for the stress pattern, consonant cluster and long vowel variables for Swedes 1 and Swedes 2, in order to determine the differential effects of altering the respective phonetic features. The scores are presented in the section regarding answering the second research question. Tables 1 and 2 suggest

Table 2. Means, standard deviations and 20% trimmed means of reaction times for all answers.

<table>
<thead>
<tr>
<th></th>
<th>Participants</th>
<th>Mean (seconds)</th>
<th>Standard deviation</th>
<th>20% trimmed mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>All sentences</td>
<td>Malians</td>
<td>1.20</td>
<td>0.67</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>Swedes 1</td>
<td>1.70</td>
<td>0.65</td>
<td>1.54</td>
</tr>
<tr>
<td></td>
<td>Swedes 2</td>
<td>1.36</td>
<td>0.61</td>
<td>1.25</td>
</tr>
<tr>
<td>Stress pattern sentences</td>
<td>Malians</td>
<td>1.29</td>
<td>1.16</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>Swedes 1</td>
<td>1.74</td>
<td>0.81</td>
<td>1.62</td>
</tr>
<tr>
<td></td>
<td>Swedes 2</td>
<td>1.20</td>
<td>0.55</td>
<td>1.13</td>
</tr>
<tr>
<td>Consonant cluster sentences</td>
<td>Malians</td>
<td>1.23</td>
<td>0.71</td>
<td>1.05</td>
</tr>
<tr>
<td></td>
<td>Swedes 1</td>
<td>2.36</td>
<td>1.24</td>
<td>2.05</td>
</tr>
<tr>
<td></td>
<td>Swedes 2</td>
<td>1.73</td>
<td>1.36</td>
<td>1.44</td>
</tr>
<tr>
<td>Long vowel sentences</td>
<td>Malians</td>
<td>1.30</td>
<td>0.78</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td>Swedes 1</td>
<td>1.60</td>
<td>0.59</td>
<td>1.45</td>
</tr>
<tr>
<td></td>
<td>Swedes 2</td>
<td>1.49</td>
<td>0.65</td>
<td>1.38</td>
</tr>
</tbody>
</table>

Note: aNumber of participants: Malaysians = 38, Swedes 1 = 30, Swedes 2 = 21. bRobust inferential t tests were performed with 20% trimmed means.

Figure 2. The 20% trimmed means of reaction times for all answers.
that Swedes 2 understood the Malaysian speaker better than Swedes 1 did, and therefore somehow more balanced mutual intelligibility appeared to be achieved as the result of altering the three phonetic features. In the following sections, we report whether this observation has statistical significance.

**Do alterations of word stress pattern, consonant clusters and vowel length increase the intelligibility of Swedish listeners’ perception of Malaysian English?**

Three sets of robust $t$ tests show that the difference between Swedes 1 and Swedes 2 in understanding the Malaysian speaker is statistically significant (see Table 3). We also ran non-parametric independent-sample Mann–Whitney $U$ tests in SPSS to double-check the significance. All of them indicated that the group difference is significant, rejecting the null hypothesis: the distribution of the variable is the same across categories of participants, confirming the test results. What is not shown in the statistical results is the effect of altering the realization of final consonants of the multisyllabic words aiming to test word stress: ‘elephant’, ‘instrument’ and ‘smallest’. It was, however, evident in the interview data that hearing all of the consonants in the multisyllabic words increased the intelligibility of the Malaysian speaker. Contrary to Swedes 1 mentioning that not hearing the final consonants of the words made it difficult to understand the speaker, Swedes 2 listeners hardly pointed them out as a source of difficulty.

The same trend is also exhibited in the robust $t$-test results for comparing the reaction times of answers for the stress pattern and consonant cluster, as seen in Table 4. However, the difference in the reaction times for responding to long vowel sentences was not statistically significant. The $t$ tests for reaction times overall seemed to suggest that Malaysian English became more intelligible to Swedish listeners when the speaker altered the three phonetic features. We did not, however, expect reaction times to be a strong indicator for the group difference. We did not urge the listener

---

**Table 3. Robust $t$-test results between Swedes 1 ($n = 30$) and Swedes 2 ($n = 21$) for answers given with understanding.**

<table>
<thead>
<tr>
<th></th>
<th>Trimmed mean difference (%)</th>
<th>$t$</th>
<th>$df$</th>
<th>$p$ value$^a$</th>
<th>Cohen’s $d^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>All sentences</td>
<td>−27.38</td>
<td>6.70</td>
<td>25.62</td>
<td>0</td>
<td>1.83</td>
</tr>
<tr>
<td>Stress pattern sentences</td>
<td>−42.22</td>
<td>8.29</td>
<td>17.00</td>
<td>0</td>
<td>2.13</td>
</tr>
<tr>
<td>Consonant cluster sentences</td>
<td>−25.11</td>
<td>4.20</td>
<td>26.10</td>
<td>0.0003</td>
<td>1.28</td>
</tr>
<tr>
<td>Long vowel sentences</td>
<td>−23.16</td>
<td>4.98</td>
<td>26.82</td>
<td>0.00003</td>
<td>1.50</td>
</tr>
</tbody>
</table>

Note: $^a$Significance level: 0.05. $^b$Effect size (small, $0.2 \leq d < 0.5$; medium, $0.5 \leq d < 0.8$; large, $d \leq 0.8$) is provided when $p < 0.05$.

**Table 4. Robust $t$-test results between Swedes 1 ($n = 30$) and Swedes 2 ($n = 21$) for reaction times.**

<table>
<thead>
<tr>
<th></th>
<th>Trimmed mean difference (seconds)</th>
<th>$t$</th>
<th>$df$</th>
<th>$p$ value$^a$</th>
<th>Cohen’s $d^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>All sentences</td>
<td>0.31</td>
<td>2.14</td>
<td>25.2</td>
<td>0.04</td>
<td>0.66</td>
</tr>
<tr>
<td>Stress pattern sentences</td>
<td>0.48</td>
<td>2.70</td>
<td>24.85</td>
<td>0.01</td>
<td>0.83</td>
</tr>
<tr>
<td>Consonant cluster sentences</td>
<td>0.63</td>
<td>2.71</td>
<td>26.71</td>
<td>0.01</td>
<td>0.63</td>
</tr>
<tr>
<td>Long vowel sentences</td>
<td>0.06</td>
<td>0.36</td>
<td>16.86</td>
<td>0.72</td>
<td>–</td>
</tr>
</tbody>
</table>

Note: $^a$Significance level: 0.05. $^b$Effect size (small, $0.2 \leq d < 0.5$; medium, $0.5 \leq d < 0.8$; large, $d \leq 0.8$) is provided when $p < 0.05$. 
participants to answer as quickly as they could, and therefore the obtained measures may be a more direct representation of subjects’ baseline behaviour.

**Do the altered realizations of the three phonetic features contribute equally to the intelligibility of Malaysian English against Swedish listeners, or do they show differential effects?**

As discussed earlier in the Method section, the data did not have a normal distribution and did not meet the assumptions for using standard parametric tests (A. Field, 2018). Therefore, for determining whether the alterations of stress pattern, consonant cluster and long vowels have made differential effects on the intelligibility of Malaysian English against Swedish listeners, a non-parametric robust multivariate analysis of variance (MANOVA) (Conover, 1999) was performed in SPSS. For this analysis, we first transformed all variables to standard scores (z values), as presented in **Table 5**.

Aligned with the robust t-test results reported previously, the result of the MANOVA first shows that there was a significant difference between Swedes 1 and Swedes 2 in understanding the Malaysian speaker, $F(2, 47) = 16.94, p = 0$; Wilk’s $\Lambda = 0.48$, partial $\eta^2 = 0.52$, and observed power was 1. More importantly, the test statistics, effect sizes and observed powers of the tests of between-subject effects presented in **Table 6** determine that, although all of them had clear effects on the intelligibility of the Malaysian English for Swedish listeners, altering the stress pattern had the largest effect; long vowels had the second largest; and the consonant cluster had the least effect. This ranking of altering effects of the three phonetic features was already implied in the test statistics, $p$ values and effect sizes of the three sets of t tests comparing the two groups of Swedish listeners (see **Table 3**).

We take the result of the MANOVA with caution, viewing that the different effects of altering the three phonetic features are rather inconclusive and require further scrutiny,

**Table 5.** Standard scores (z values) of answers given with understanding by Swedes 1 and Swedes 2.

<table>
<thead>
<tr>
<th>Participants</th>
<th>Stress pattern</th>
<th>Consonant cluster</th>
<th>Long vowel</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean Standard deviation</td>
<td>Mean Standard deviation</td>
<td>Mean Standard deviation</td>
</tr>
<tr>
<td>Swedes 1</td>
<td>$-0.5794$ 0.90680</td>
<td>$-0.4116$ 0.86282</td>
<td>$-0.4619$ 0.95688</td>
</tr>
<tr>
<td>Swedes 2</td>
<td>$8277$ 0.29140</td>
<td>$5880$ 0.89629</td>
<td>$0.6599$ 0.62851</td>
</tr>
</tbody>
</table>

Note: “Number of participants: Swedes 1 = 30, Swedes 2 = 21.

**Table 6.** Differential effects of altering the three phonetic features on intelligibility of Malaysian English for Swedes 1 and Swedes 2.

<table>
<thead>
<tr>
<th>Source</th>
<th>Variable</th>
<th>df</th>
<th>Mean square</th>
<th>F</th>
<th>p value</th>
<th>Partial $\eta^2$</th>
<th>Observed power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>Stress pattern</td>
<td>1</td>
<td>24.46</td>
<td>46.91</td>
<td>0</td>
<td>0.49</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Consonant cluster</td>
<td>1</td>
<td>12.34</td>
<td>16.06</td>
<td>0</td>
<td>0.25</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td>Long vowel</td>
<td>1</td>
<td>15.55</td>
<td>22.11</td>
<td>0</td>
<td>0.31</td>
<td>0.96</td>
</tr>
<tr>
<td>Error</td>
<td>Stress pattern</td>
<td>49</td>
<td>0.52</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Consonant cluster</td>
<td>49</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Long vowel</td>
<td>49</td>
<td>0.70</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: “Effect size of partial $\eta^2$ can be understood as small ($0.01 \leq \eta^2 < 0.06$), medium ($0.06 \leq \eta^2 < 0.14$) or large ($\eta^2 \leq 0.14$).
mainly due to the following two reasons. First, although they exhibit difference in sizes, it should be noted that all three effects are very strong and large. Second, as mentioned in the Method section, the Malaysian speaker happened to change the quality of some vowels in the consonant clusters words, while trying to realize consonant clusters clearly. This in fact seemed to affect the listeners’ perception somewhat negatively. For example, some listeners mentioned that they heard the consonant cluster word ‘kids’ as something like ‘kads’ as the speaker tried to realize the /d/ clearly.

**Do alterations of the three phonetic features result in balanced, mutual intelligibility between Malaysian and Swedish English?**

As Jeong et al. (2017) report, Malaysian listeners understood the Swedish speaker well. The question that we pose here is whether a balanced mutual intelligibility was achieved between Swedish English and Malaysian English when the three phonetic features of Malaysian English were altered. Given the large discrepancy in the comparison between Malaysian listeners (see Jeong et al., 2017), overall, Tables 7 and 8 show that the three altered phonetic features of Malaysian English have led Malaysians and Swedes to obtain results that are more similar to each other than those of the first experiment.

There is no statistical difference in reaction times for all three types of sentences, as well as in the answers given, with understanding of stress pattern sentences. However, Swedes 2 still struggled to understand the Malaysian speaker’s consonant clusters, even after the alteration, more than Malaysians listening to the Swedish speaker, although they were able to understand her much better than Swedes 1 understood the same Malaysian speaker. As mentioned previously, this seems due to the Malaysian speaker having been less successful in altering her ways of pronouncing consonant clusters than long vowels and stress patterns. What is interesting is that her altered long vowel sounds reversed the degrees of intelligibility of Malaysian English and Swedish English against Swedes and Malaysians: Swedes 2 listeners performed better than Malaysian listeners for understanding the long vowel sentences.

| Table 7. Robust t-test results between Malaysians and Swedes 2 for answers given with understanding. |
|-------------------------------------------------|--------------------|-----------------|----------------|------------------|---|
| Trimming mean difference (%) | t | df | p value<sup>a</sup> | Cohen’s d<sup>b</sup> | 
| All sentences | 2.68 | 1.09 | 24.67 | 0.29 | – |
| Stress pattern sentences | –5 | 2.01 | 23.00 | 0.06 | – |
| Consonant cluster sentences | 25.24 | 4.53 | 24.07 | 0.0001 | 1.29 |
| Long vowel sentences | –11.22 | 2.65 | 23.83 | 0.01 | 1.50 |

Note: <sup>a</sup>Significance level: 0.05. <sup>b</sup>Effect size (small, 0.2 ≤ d < 0.5; medium, 0.5 ≤ d < 0.8; large, d ≤ 0.8) is provided when p < 0.05.

| Table 8. Robust t-test results between Malaysians and Swedes 2 for reaction times. |
|-------------------------------------------------|--------------------|-----------------|----------------|------------------|---|
| Trimming mean difference (seconds) | t | df | p value<sup>a</sup> | Cohen’s d<sup>b</sup> | 
| All sentences | –0.21 | 1.58 | 22.55 | 0.13 | – |
| Stress pattern sentences | –0.15 | 0.97 | 19.24 | 0.34 | – |
| Consonant cluster sentences | –0.36 | 1.89 | 19.6 | 0.07 | – |
| Long vowel sentences | –0.25 | 1.37 | 17.34 | 0.19 | – |

Note: <sup>a</sup>Significance level: 0.05. <sup>b</sup>Effect size (small, 0.2 ≤ d < 0.5; medium, 0.5 ≤ d < 0.8; large, d ≤ 0.8) is provided when p < 0.05.
Discussion and conclusion

This study has found the following:

1. Alteration of the three phonetic features (word stress pattern, consonant cluster and long vowel) had strong, positive effects on the intelligibility of Malaysian English for Swedish listeners.

2. When compared to the degrees of the alteration effect, altering stress patterns appeared to have the strongest effect, lengthening vowels the second strongest and consonant cluster articulation the least.

3. Alteration of the three phonetic features in Malaysian English resulted in more balanced results obtained from statistical analyses between Malaysian and Swedish English than before.

The second finding is more tentative than the other two. First, this is because the Malaysian speaker was not completely successful at altering some target consonant clusters. Second, some consonant cluster test items that clearly improved the intelligibility were in stress pattern and long vowel testing sentences. In addition, it can be discussed how well stimuli comprised of the readings of one speaker represent a speech community, although using a single speaker for perception tests is found in the literature (e.g., Tajima et al., 1996), and as reported in the Method section, the speaker in our study had generic phonetic features of Malaysian English.

Despite these issues, the results overall show that how a speaker pronounces the three phonetic features we tested can be crucial for using ELF. Our study firstly confirms Jenkins’ (2000, 2002, 2015) syllabus of the Lingua Franca Phonetic Core, in relation to how consonant clusters and long vowels in words need to be realized. They also support the previous studies which report, quite often, that each phone in the consonant cluster is important information for the listener and, thus, needs to be pronounced (Jenkins, 2002; Lesley, 2014; Suenobu et al., 1992; Tajima et al., 1996). As for the vowel length, Derwing and Munro (2015) argue that what distinguishes an English tense vowel from a lax one is not duration but quality difference. Although the long–short vowel distinction may involve quality change (i.e. tense or lax), our data suggest that the duration change in vowel length certainly has an effect on the speaker’s intelligibility, aligning with Bent et al. (2008), Quené and van Delft (2010), Smith et al. (2003) and Tajima et al. (1996).

What makes our study significant is the finding that suggests revising the LFC syllabus to include word stress patterns. Jenkins has regarded word stress patterns as a non-core feature, reasoning that it ‘varies considerably even across L1 Englishes’ (Jenkins, 2015, p. 92). However, although stress patterns for some words are more diverse than others across different varieties of English, we question whether stress patterns of many frequent multisyllabic words in ELF actually vary and what different stress placements of such words cause. The effect of placing word stress correctly may have been enlarged, because the Swedish rely on the stress pattern as an important acoustic cue, perhaps more than some other linguistic groups. However, J. Field (2005) and Kashiwagi and Snyder (2008) found that listeners whose first languages do not use word stress as an acoustic cue still struggled when stress was placed unexpectedly. Both
their study and ours together suggest the word stress pattern as a core feature, regardless of whether the interlocutor is a native speaker or a non-native speaker.

Here, we wish to extend our discussion of the relevance of our study to using English in international contexts. Many of the previous studies on intelligible English pronunciation used native listeners for judging non-native speakers. Unlike these studies, we have suggested that the ways that the Malaysian speaker realized stress patterns, long vowels and consonant clusters after she altered them can be regarded as phonetic core features of ELF. We find the ground for this suggestion from Swedish English being a similect (Mauranen, 2012) and Malaysian English being a dialect of English.

An English dialect, whether it is one of the Inner Circle or Expanding Circle Englishes, has its own unique syntactic, morphologic, stylistic and phonetic features shared within the speech community where it has developed (Kachru, 1992). On the contrary, according to Mauranen (2012), speakers of a similect of English do not have their own English-using community. They share the L1, and therefore, their English similect has features shaped through the influence of their L1, which is similar to non-native English varieties of English. However, since the speakers of an English similect use it to communicate international interlocutors, it does not develop new linguistic features in the same way a dialect does, through the interaction of its users. A similect does ‘not undergo sound changes [and] accent diversification’ that are shared only within a speech community, but it is renewable under cross-linguistic influence (Mauranen, 2012, p. 29).

Descriptions of Malaysian English and Swedish English in the literature clearly indicate that the former is a dialect, while the latter, like many other European Englishes, is a similect (Björkman, 2014; Low, 2010; Seidlhofer, 2010). Therefore, first, we view that the localized stress pattern, consonant cluster and long vowel in Malaysian English as a dialect were not easy for Swedish listeners who are not members of the Malaysian English community. The alteration of the three features may have given Malaysian English a lingua franca status, resulting in much higher intelligibility of the Malaysian speaker than before the alteration. Jenkins (2015) suggests that, for native and non-native speakers alike, being able to accommodate to the interlocutor is the key to successful communication in ELF contexts. Understanding Malaysian English as a dialect and the alteration of some of its phonetic features as what led to improving its intelligibility can be relevant to what Jenkins means by ability to accommodate or adaptability as important part of international communication skills.

On the other hand, the easiness of the Swedish speaker’s accent felt by Malaysian listeners may be that Swedish English, as a similect, has phonetic characteristics that are intelligible in international communication. For example, the Swedish speaker, while speaking Swedish-accented English, clearly realized the core features in the LFC and stress pattern in accordance with Inner Circle Englishes. We are aware that the pronunciations of some similect varieties can be less intelligible than others. However, it needs to be noted that here we are concerned with proficient speakers of similects and dialects rather than with those with low proficiency that often have been the subjects of previous studies on intelligibility of non-native speakers.

As a pedagogical implication of our study, we wish to discuss the two aspects of pronunciation and intelligibility that the users of ELF need to relate to as both listeners and speakers. As listeners, it is impossible for us to expect that our international
interlocutors always have intelligible pronunciation. Thus, it is desirable to expose learners to diverse dialects and similects of English so that their perception can be trained to understand a wide range of phonological variations. In addition, as speakers, ELF users need to be responsible for enabling others to understand them well. As Pillai (2017) argues, simply being a speaker of one certain dialect of English does not necessarily make the person intelligible or unintelligible. However, as she also emphasizes, when talking to international interlocutors, ELF speakers may need to make efforts to alter certain core features that are likely to confuse non-local listeners while maintaining other features of their local accents. The points that Pillai makes are with reference to the context of English in Malaysia. However, they can be also very relevant to other Asian Englishes, particularly to those regarded as dialects. English teachers in the countries where people have developed their own English dialects can encourage such an attitude among students, together with teaching phonetic features important for intelligibility in international contexts.

Acknowledgements

The authors want to express gratitude to the three reviewers and the editor, Leah Gilner, as their comments have been invaluable.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Hyeseung Jeong http://orcid.org/0000-0003-1058-7637
Bosse Thorén http://orcid.org/0000-0001-7966-320X
Juliana Othman http://orcid.org/0000-0001-5151-099X

References


