Using discuss-the-differences tasks to evaluate focused intelligibility in English as a lingua franca research

Christine Lewis

Universiti Brunei Darussalam

Abstract

As the utilization of English as a lingua franca (ELF) continues to increase, so does the need for additional research on the intelligibility of ELF exchanges. Most ELF research is currently based on naturally occurring data, but researchers might also consider evaluating the intelligibility of pronunciation with focused tasks as well. A discuss-the-differences task is one way of measuring focused ELF intelligibility as it incorporates what might be described as three types of research methods: ELF, pronunciation, and intelligibility. In Brunei Darussalam, a set of discuss-the-differences tasks was designed to explore the use of word stress in ELF interactions between ASEAN speakers. This article explores why this type of task was chosen, how it was implemented in Brunei, and some of the preliminary results from this new corpus.

Introduction

Nowadays, about one-third of the world's population can communicate in English, and less than a quarter of those speakers are English as a native language (ENL) speakers using English as their first language (L1) (Crystal, 2012). In fact, ENL speakers are not only considered the minority of English users now, but are also not involved in the majority of the English interactions that occur (Seidlhofer, 2001). These non-ENL focused exchanges are considered English as a lingua franca (ELF) interactions. Jenkins (2015) defines ELF speakers as "those who use English primarily with non-native English speakers from other L1's than their own rather than primarily with [ENL] speakers" (p. 11).

The increasingly accepted ELF approach to English calls for a shift from the past *nativeness* focus, in which speakers were expected to imitate an ENL accent, to an *intelligibility* focus, in which ELF speakers' primary goal is to be understood (Jenkins, 2007; Levis, 2005; Seidlhofer, 2011). Intelligibility refers to "the extent to which a speaker's message is understood by a listener" (Munro & Derwing, 1999, p. 379). Unfortunately, most pronunciation research continues to encourage ENL standards (Thomson & Derwing, 2015) and intelligibility research has often utilized ENL speakers as the main evaluators (Rajadurai, 2007). With the increase of ELF exchanges, there is a need for updated intelligibility research which includes a variety of ELF speakers as listeners and judges of what is understandable.

In some of the initial research on ELF intelligibility, it has been reported that pronunciation plays a major role in contributing to misunderstandings (Deterding, 2013; Jenkins, 2000) and by examining which pronunciation features impede intelligibility in ELF exchanges, it is possible to gain a heightened understanding of which ones are most beneficial for ELF speakers to acquire (Kirkpatrick, 2010). It may be that these essential pronunciation features differ by region (Deterding & Kirkpatrick, 2006). Nevertheless, Jenkins (2000) has suggested a Lingua Franca Core (LFC) of those features which she believes enhance intelligibility. Much of the original data on the intelligibility of ELF was derived from naturally occurring speech; and while it is important to examine this kind of speech, it is also crucial to consider more focused speech as learner avoidance strategies may sometimes obscure features that cause difficulties for speakers (Derwing & Munro, 2005). Thus, as additional ELF intelligibility research occurs, focused ELF intelligibility research should be considered along with naturally occurring data.

Discuss-the-differences tasks, which combine what might be described as ELF, pronunciation, and intelligibility research methods, have been used to gather spontaneous ELF

data focused on word stress in polysyllabic words. This article reviews the research and design behind the discuss-the-differences tasks, the implementation of them in Brunei Darussalam, and some of the initial findings from that corpus.

Research behind the method

One method of gathering data for focused ELF intelligibility research is the use of discuss-the-differences tasks. In general, speakers discuss two pictures (picture A and picture B) which have a number of similarities as well as some variations. One participant views picture A while the other participant from a differing L1 background refers to picture B. Then without observing each other's drawings, they talk about how they differ. This activity is slightly different than a 'find/spot-the-differences' task as participants are encouraged to describe variances in more detail than merely stating 'Mine is different'. Ideally, participants explain in detail what makes their drawings distinct.

These tasks integrate three types of research paradigms that could be labeled as ELF, pronunciation, and intelligibility research methods, each of which will be outlined below. While most ELF exchanges focus on intelligibility, not all intelligibility research includes the ELF paradigm; thus, these two research methods will be considered as separate for this article. The combination of these methods ends up with a task that encourages interactivity with immediate feedback from ELF listeners, fosters an environment for the production of specific pronunciation features, minimizes the need for top-down listening strategies, and offers the researcher contextualized language.

ELF research

ELF research has often been based on corpora such as the Vienna-Oxford International Corpus of English (VOICE, 2013) which is comprised of non-scripted, interactive, naturally occurring interchanges between speakers of English from various L1 backgrounds. With intelligibility as the focus of ELF exchanges, reception of the language is just as important as production and "the interactivity criterion allows [researchers] to investigate how people react to and experience ELF" (Breiteneder, Pitzl, Majewski, & Klimpfinger, 2006, p. 168). The interactivity requirement means that ideally everyone who is speaking is also a listener and vice versa. This method encourages listeners with various L1 backgrounds to be judges of what is intelligible.

When considering activities that are most interactive, research promotes a two-way, closed-solution task where some planning is allowed and which also requires a high degree of negotiating for meaning (Folse, 2006; Long, 1989). A discuss-the-differences task creates this environment for non-scripted interactivity by encouraging participants to listen to each other and respond with how their own drawing compares to their partner's. As there is a closed solution – an appropriate or inappropriate response to the speaker's utterances – negotiation can take place. For example, if one speaker states that they have a calendar on the wall in their picture, the other speaker knows they should react to that statement while referring to their own drawing. Either they also have a calendar on the wall, or they do not. If they do not, they should ideally state how their picture differs (maybe their calendar is on the table). Usually the listener's reactions to each comment allow the researcher to identify where a speaker may have been misunderstood, and therefore it is the listener who is the evaluator of what is intelligible and not the researcher. In summary, the discuss-the-differences task, when used with ELF speakers, fulfills most of the VOICE standards for ELF data. However, the task purposely violates the VOICE criteria that interactions must occur in natural settings because analyzing the use of specific language features can also be beneficial in ELF research.

Pronunciation research

In pronunciation research, it is not ideal to exclusively use data that occurs in natural situations due to the fact that speakers often employ learner avoidance. Derwing and Munro (2005) remind researchers that "observing only errors that occur in natural productions may conceal underlying processes because of learner avoidance strategies" (p. 381). Since speakers may not attempt more difficult language, naturally occurring data might not represent a full range. Learner avoidance is usually an advantageous strategy for learning languages and can be especially valuable in ELF interactions, but researchers should take this into consideration when studying conversations that transpire naturally and explore additional types of data as well.

When strategically designed, a discuss-the-differences task can help circumvent learner avoidance. The researcher can choose the lexis they wish to hear participants utilize and then design their drawings accordingly. As the impact of innovative word stress on the intelligibility of ELF conversations is still debated (Cruttenden, 2014; Deterding, 2013; Jenkins, 2000; Lewis & Deterding, 2018), the drawings in the current research were designed so that participants would produce a variety of polysyllabic words to be analyzed later. The development of the set of discuss-the-differences tasks will be described in more detail following a discussion about the remaining type of research contained in these tasks.

Intelligibility research

The third type of research method the discuss-the-differences task incorporates is intelligibility research. Some common methods of assessing intelligibility are based on functional testing, evaluating how thoroughly listeners understand the speakers' intended message (Kang, Thomson, & Moran, 2018). These types of tasks, such as transcriptions, cloze tests, and true/false (T/F) judgements, allow the researcher to assess the accuracy of what listeners hear on a phonemic level. For transcriptions, listeners write every word they hear in the utterance; for cloze activities, some of the words are already provided and listeners fill in the blanks of the missing key words; and in T/F assessments, listeners decide whether the content of the utterance is true or not (e.g., if they hear 'A bird has four legs', the listener ideally would answer that this is a false statement).

Kang et al. (2018) evaluated these intelligibility tests, in addition to some others, in an effort to discover which tasks best predict how well speakers might be understood on longer listening comprehension texts. One of the additional tasks they included was the transcription of nonsense sentences which were semantically illogical but followed standard syntax (e.g., 'A shy chair can smell the fat train'). Because a lack of background knowledge on a topic can adversely affect the intelligibility of speakers and muddy research results (Picheny, Durlach, & Braida, 1985), a nonsense task is beneficial given that it removes the need for context and might shift the focus more onto the pronunciation.

After evaluating T/F statement, nonsense sentence, transcription, filtered sentence (cloze activity), and scalar rating (listener judgment) tasks, Kang et al. (2018) ascertained that not every assessment evaluates all pronunciation features equally. For example, the task involving nonsense sentences assesses almost all phonological variables except for prosody-enriching features such as word stress. Prosody was instead found to be integral for intelligibility in T/F sentence tasks. Consequently, researchers should consider what pronunciation features they want to test when choosing specific intelligibility tasks. For determining how well someone would be understood on a longer listening comprehension task, Kang et al. (2018) concluded that the most precise way of gauging intelligibility was through the nonsense sentence tasks even though they do not assess suprasegmental features, and that T/F sentence tasks, which do evaluate prosody, were the second-best predictor of intelligibility.

As a type of enhanced T/F task, a discuss-the-differences task encourages the listener to decide if the speaker is describing something that is the same (true) as the images in their

partner's picture or different (false). However, instead of the task being a one-sided listening task with no background knowledge, a discuss-the-differences task provides opportunities for the crucial component of interactivity and also minimizes the amount of background knowledge needed. Just as Kang et al. (2018) observed, intelligibility is easier to evaluate when listeners do not need additional information to comprehend utterances.

Jenkins (2000) agrees that the listeners' background knowledge plays an important role in the intelligibility of ELF English. She argues that "even at relatively high levels of competence, [second language English listeners] still process speech using a predominance of bottom-up strategies" (p. 80). They tend to concentrate more on the smaller units of meaning (i.e., sounds and words) in an utterance and they make less reference to their own personal knowledge, experience, and/or context (top-down processing). It is especially difficult for ELF communicators to use top-down strategies since they do not have a shared cultural background with their interlocutors. Rajadurai (2007) also encourages the consideration of context in intelligibility research and argues that many of the previous studies "ignore the fact that speech is context-specific and highly dependent on the topic, participants, and situation" (p. 90).

In the original research on the LFC by Jenkins (2000), some of her data was derived from students describing pictures to each other. Each student had the same set of pictures and they took turns describing one of the pictures while the listener would point and guess which picture was being described. Misunderstandings continued to occur despite the fact that they had a clear context in front of them. However, had there not been a shared framework, the assumption is that there may have been more misunderstandings. Thus, using pictures is one way to decrease the demand of background knowledge in intelligibility research.

Since participants have mostly similar pictures in front of them in a discuss-the-differences task, the need for background knowledge is reduced and participants can focus on communication. The participants share the same general understanding as they begin speaking, and emphasis is on the pronunciation of what is communicated.

Finally, intelligibility research requires investigators to know what speakers intended to say (Munro & Derwing, 2015). In one-way listening tasks, speakers are either given a script to read or they are asked what they intended to say before their recordings are evaluated by others. However, in an interactive exchange, with somewhat spontaneous conversation, knowing the intentions of the speaker can be more challenging. Therefore, ideally a chosen task should facilitate identification of intended utterances and minimize the need for follow-up. Inquiring what speakers aimed to say can create some anxiety for them, so it is best to reduce the need for clarification. A discuss-the-differences task can reduce the need to follow-up on speaker utterances since the drawings encourage expected vocabulary and make it easier for the researcher to predict speaker intentions.

Method summary

A discuss-the-differences task accomplishes a combination of ELF, pronunciation, and intelligibility research methods. It creates an environment for:

- non-scripted interaction of language
- participants as the judges of the language (with true/false reactions)
- encouragement of focused pronunciation features (e.g., polysyllabic words)
- low demand of background knowledge using simple pictures
- contextualized utterances

Interactivity occurs as the partners alternate discussing their pictures. Listeners are also obliged to agree or disagree with what they hear (T/F statements), making them the evaluators of each other's pronunciation. The objects in the different drawings encourage production of specific pronunciation features and also provide a mutual framework for participants which

minimizes the amount of background knowledge needed. Finally, the drawings often allow researchers to predict what the speakers intended to say and reduce the amount of follow-up needed for the speakers.

A polysyllabic design

In his resource book for practicing pronunciation, Hancock (1995, pp. 86–89) includes a find-the-differences task devised to practice the stress of compound nouns. His task also adopts a possible storyline between the two pictures. Inspired by that activity, the goal of the current research was to design two pairs of drawings with a storyline between them that would encourage participants to produce a variety of polysyllabic words for analysis of the impact of word stress variation on intelligibility among speakers in ASEAN.

The researcher initially compiled a list of polysyllabic words she hoped the participants might produce and then tried to imagine environments that would incorporate the majority of them. Eventually a living room scene and a street scene were chosen. Then, a local artist, Richard Chin, was hired to illustrate two drawings for each scene. He understood the vision and purpose behind the drawings and was able to suggest alternate ways of including the desired lexical items. Two pairs of pictures were eventually designed in a way that would encourage interlocutors to produce a variety of polysyllabic words.

The objects chosen for the pictures are based on an intermediate speaker's lexical knowledge, though a few items for advanced speakers are included (i.e., *aquarium, pacifier/dummy, baton*). The objects are in differing situations in the two pictures in the hope that a variety of words, not just nouns, will be spoken by the participants (i.e., *whistling, surprised, under*). The drawings also include a few written words such as *conference, departure, Japanese*, and *electronics* that participants might read.

The pictures were devised in a way that would require the participants to produce lexical items that are commonly studied in a pronunciation class. Each set of pictures displays times and other numbers which are often confused (i.e., *fourTEEN* vs. *FORty, fifTEEN* vs. *FIFty*). Customarily, pronunciation lessons present a basic rule that, for disyllabic words, stress tends to be produced on the initial syllable in nouns and the second syllable in verbs, but there are exceptions. Thus, the drawings include objects that are not pronounced according to those rules such as *balLOON, giRAFFE,* and *guiTAR*. Some situations in the drawings also encourage interlocutors to produce adjectives (i.e., *CIRcled, INjured, MEssy*). Finally, certain scenes in the pictures provide the opportunity for the speakers to use words that have shifted stress depending on their suffixes: *graduATion, phoTOgrapher,* and *muSIcians*.

The two sets of pictures each have one 'organized' picture where life seems to be going on as normal and one 'chaotic' picture where everything has changed for the worse. The pair of pictures that take place in a living room involve a family of five including a mother who is preparing to travel and the drawings are labeled as either Living Room Organized (LO) or Living Room Chaotic (LC). The street scenes show a busy street in a city where robbers are targeting a jewelry store and those pictures are labeled as Street Organized (SO) or Street Chaotic (SC). Both sets are entertaining to look at and designed to create an atmosphere where speakers forget they are being recorded and are curious to know how their partner's picture differs. Figures 1 to 4 show the final drawings at half the size.
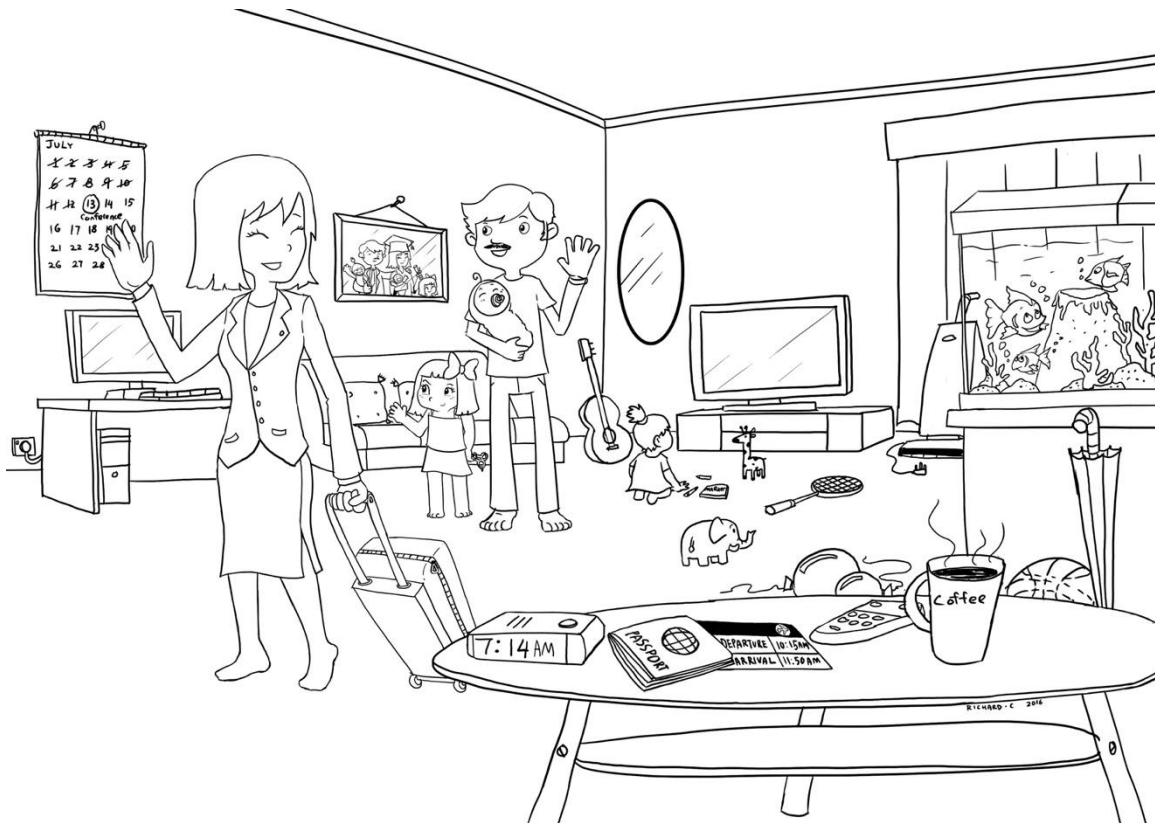
Figure 1. Living Room Organized (LO), Set 1, Picture A

Figure 2. Living Room Chaotic (LC), Set 1, Picture B
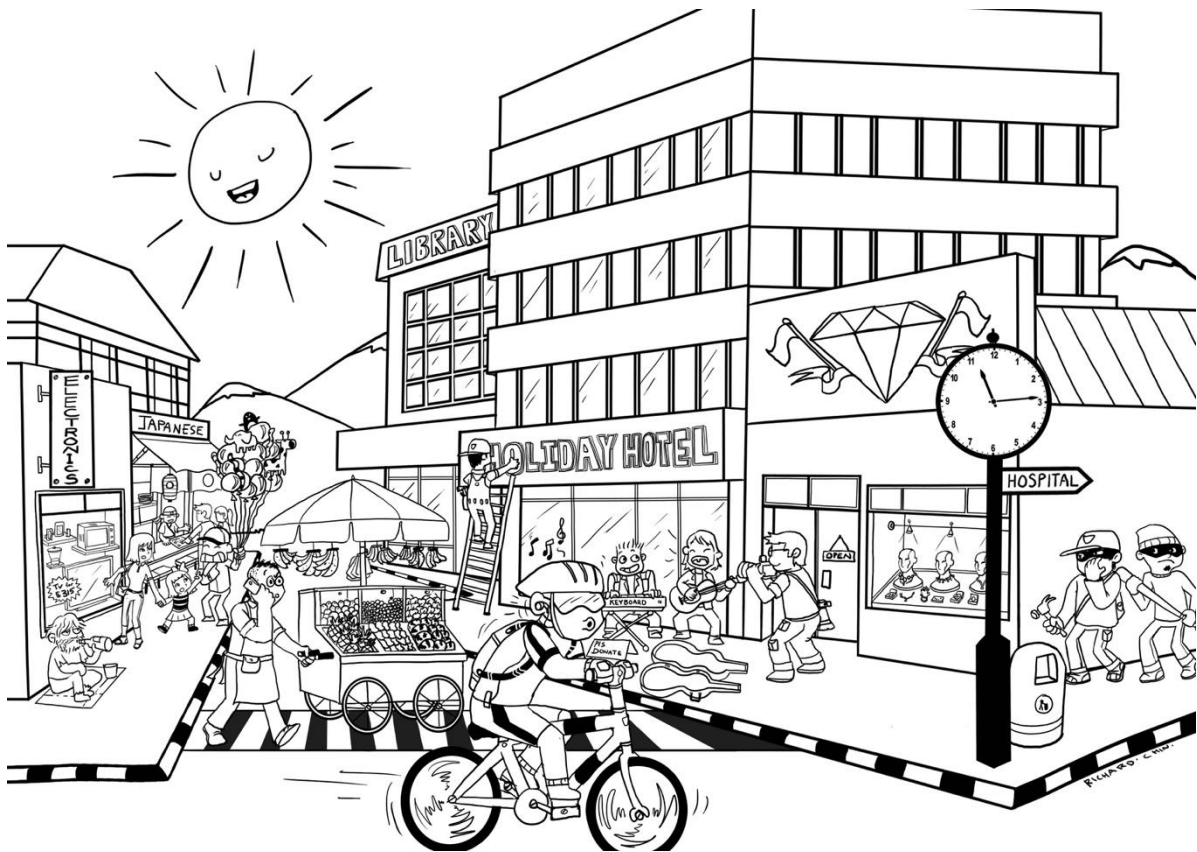
Figure 3. Street Organized (SO), Set 2, Picture A

Figure 4. Street Chaotic (SC), Set 2, Picture B

Implementation

These discuss-the-differences tasks designed to encourage the production of polysyllabic words were implemented in Southeast Asia. Over 40 speakers from countries from the Association of Southeast Asian Nations (ASEAN) region participated in the two tasks and about 90% of the participants completed follow-up surveys.

Participants

An English-enrichment project comprised of participants from all ten ASEAN countries took place at Universiti Brunei Darussalam. When offered the opportunity to participate in a study focusing on English pronunciation in the ASEAN community, 41 members of the project volunteered. The breakdown in Table 1 lists the 26 ASEAN females and 15 males that were recorded, representing nine of the ten ASEAN countries.

No.	Country	Females	Males	Total
1	Brunei (Bn)	1	1	2
2	Cambodia (Cb)	1	4	5
3	Indonesia (In)	4	1	5
4	Laos (La)	4	4	8
5	Malaysia (Ma)	3	0	3
6	Myanmar (Mm)	2	2	4
7	Philippines (Ph)	2	0	2
8	Thailand (Th)	4	2	6
9	Vietnam (Vn)	5	1	6
	Total	**26**	**15**	**41**

Table 1. Description of research participants

Participants' ages were between 26 and 44 years old with 70% of them being in their 30's. For admission to the project, each participant had been required to score at least a 5.5 on the IELTS or equivalent; and each one also took a pre-test on arrival for classroom placement. The results of the project's English placement exam placed 23 participants at an intermediate level and 18 at an advanced level.

The participants spoke a variety of first languages: Bahasa Selayar, Burmese, English, Filipino, French, Hmong, Indonesian, Javanese, Kapampangan, Khmer, Lao, Malay, Myanmar, Padang, Thai, and Vietnamese. Some participants had also studied other foreign languages in addition to English: Arabic, Chinese, French, Hindi, Indonesian, Japanese, Thai, and Vietnamese.

Recordings

The volunteers consented to participate in two interactive tasks with two other participants with L1s differing from theirs. An attempt was made to have volunteers speak with one interlocutor who communicated at around the same English level as themselves and one who spoke at a different level.

Participants sat diagonal to each other with about five feet between them which was enough distance to hinder their view of their partner's picture but close enough to hear each other well. A recorder was placed on the table in front of them and each participant wore a lapel microphone. One participant received the organized picture while the other one had the chaotic picture. Once settled, participants were instructed that they had similar pictures with similar people and objects. However, they also had a few differences (maybe the location of the objects, what the people were doing, or how the objects were being used by people). They

were informed that their task was not just to find differences but also to discuss *how* their drawings differed. They were specifically directed not to reply solely with 'Mine is different'.

The participants were given about 30 seconds to first examine their pictures and establish an overall awareness of the topic and vocabulary they would be discussing. Once both stated they were ready, the facilitator started the recorder and left the room. The pair had seven minutes to discuss as many differences as possible.

In order to investigate whether speech might be different when negotiation was no longer necessary to complete the task, after the first seven minutes, the researcher asked the participants to position the pictures side-by-side on the table and discuss them for an additional three minutes. With both drawings in view at the same time, the participants worked together to deduce what might have transpired between the two pictures and what might have caused the differences between them.

After the recording, the partner describing the organized picture left the room and the person with the chaotic picture remained. A new partner from a differing L1 background entered the room and the process began again. This time, however, the speaker with experience completing the task, the person who had previously been holding a chaotic picture, was now given the organized picture from the other set. Since this person already knew there was a storyline behind the pictures, there was additional motivation for negotiation of meaning given that they were curious about what changes had occurred to their picture.

The researcher collected 40 recordings from 41 ASEAN speakers. After one participant unexpectedly rescheduled to a later time, an African student filled in for the initial timeslot. Therefore, originally there were 42 recordings of 42 people, but the two recordings with the African speaker are not included in the current research.

Participant Feedback

As mentioned earlier, one key to intelligibility research is knowing what the speaker said and what the listener heard. Therefore, following the recording sessions, the researcher listened to each recording and noted clear examples of misunderstandings as well as other possible instances. There were 14 occasions in which listeners asked for clarification and five others where there were obvious misunderstandings – some of which included word stress variance. To find the less easily identifiable misunderstandings, the researcher extracted clips of over 500 instances in which the listener was either silent, laughed, changed the subject, or had an unexpected response. Those 500+ possibilities required feedback from either the listener, the speaker, or both.

Online surveys were sent to each participant for clarification on what was either said or heard in those 500+ clips. Each survey contained copies of the pictures the participant was viewing at the time of their recordings, audio clips of the potential misunderstandings, and a transcription of the audio with blanks where possible miscommunication occurred (a type of cloze activity). Sometimes participants responded to a recording of themselves, so the researcher could be certain what they had said; and at other times, they listened to a recording of their partner, so the researcher could know what they heard.

Figure 5 shows an example of the survey questions sent to each participant through SoGoSurvey (2017). Even though the researcher was almost certain that in audio clip 11g the Cambodian speaker intended to say that the number thirteen was 'circled' in his picture, the speaker's pronunciation was [səˈkɔː] and the listener did not seem to understand. Therefore, this clip was sent to both the listener and the speaker. Both participants listened to it and wrote the one word or phrase they thought they heard or said. If they could not hazard a guess, they were encouraged to write "??" in the response field.

First Picture
If you need a reminder, here is a copy of the first picture you discussed with **FId2.**

Living Room- C

11g

* 7.

11g- MCb1: And on the desktop computer (.) uh there is (.) a calendar (.) which (.) uh number thirteen (.) is _____ the conference.

Characters Remaining: 100

Figure 5. Sample of online feedback

The number of questions each participant answered varied, but in general, each one answered around eight to ten questions per picture. In the original feedback surveys, real names were used, so the listeners could remember who they had been speaking with. They have been deleted here and replaced with research identification codes (FId2 and MCb1 in Figure 5) to ensure anonymity.

Out of 41 participants, 37 completed the online feedback surveys. For various reasons, the remaining four did not submit feedback. The answers from the surveys were compiled and instances in which the speaker's and listener's answers differed were noted. For example, the Cambodian speaker in Figure 5 confirmed through survey feedback that he intended to say *circled*. In her feedback, the Indonesian listener replied that she heard 'called'. Therefore this clip was identified as a token of misunderstanding. As the speaker placed his stress unexpectedly on the second syllable (*cirCLED*), this token was also labeled as a misunderstanding influenced by innovative stress. Based on the survey results, around 200 possible tokens of misunderstanding, some of which include word stress, have been found and will be examined in more detail.

Some initial findings

As stated earlier, there were 19 clear misunderstandings in the participants' interactions; and based on participant feedback, around 200 potential additional tokens of misunderstandings have been identified. From that data, around 170 different words were involved in the misunderstandings and 120 of those words were polysyllabic, though stress is not implicated in every case.

Though this corpus is still in the process of being evaluated, there are some preliminary findings. Unexpected word stress does occur in ELF exchanges. In some cases, the variant lexical stress had no impact on the intelligibility of the utterance, but at other times, unexpected word stress seems to have interfered with understanding. The following are some of the initial findings on word stress already published in Lewis and Deterding (2018).

Example 1 demonstrates one of the 19 clear misunderstandings from this corpus. While a Vietnamese male and an Indonesian female were discussing the living room scene, the male observing the LO drawing, inquired about the balloons in his partner's picture. In line 1, he stressed *balloon* on the first syllable, which was a common variant among speakers in this corpus; and the listener had a difficult time comprehending him, even after he repeated the word twice with the same stress pattern and added *-s* as well in line 3. (Capital letters indicate where the speaker pronounced stress; <1> and <1> indicate the start and end of overlapping speech; and a question mark signals rising intonation. Short pauses are marked with (.), and in pauses of more than 0.3 seconds, the length of silence is included.)

(1) *1*   MVn   :   how about the **BALloon**?
     2   FId1 :   <1> the? </1>
     3   MVn   :   <1> that </1> i have the (.) er two BALloon (.) s (1.2) two BALloons
     4   FId1 :   balLOONS? <2> no </2>
     5   MVn   :   <2> yeah </2> (.) you don't have it?
     6   FId1 :   no   (Lewis & Deterding, 2018, p. 169)

There were also clear misunderstandings in which numbers were involved, such as Example 2. A Malaysian speaker and a Cambodian speaker had different times on the clocks in their living room pictures; however, they came to an agreement that their times were the same. The Cambodian speaker, discussing the LC picture, began explaining his time in line 4 with stress on the first syllable of *for-,* as is standard, but hesitated. When he restarted, he shifted his stress to the last syllable of *forty,* which may have contributed to the confusion. The Malaysian listener did not realize he was saying a different time than hers.

(2) *1*   FMa   :   and uh my time here written seven fourteen A.M (1.5)
     2   MCb1 :   hm-mm
     3   FMa   :   do you have the time there? (1.0)
     4   MCb1 :   seven FOR- (.) **forTY**?
     5   FMa   :   yeah seven **fourTEEN**
     6   MCb1 :   it the same? yeah?
     7   FMa   :   <1> oh just the same </1>
     8   MCb1 :   <1> it but uh the </1> time? is on the table?
                                                                    (Lewis & Deterding, 2018, p. 169)

Table 2 depicts some additional tokens of misunderstandings involving lexical stress. The table shows which participant spoke each utterance and which one was listening. The *Word(s)* column explains which word the speaker intended to say, and the *Heard* column explains what the listener understood. The responses in tokens 6 and 7 may seem surprising, but listeners were encouraged to write what they heard, even if it did not make sense. When the listener could not hazard a guess about what was said, they were asked to answered with '??'

as in tokens 4 and 5. Finally, the last column shows the context for the misunderstood word and the non-standard stress pattern that was used.

From the living room scene discussions, the word *circled* in token 3 was heard as 'called' (as mentioned in the discussion about Figure 5), the word *umbrella* in token 5 could not be deciphered, and the phrase *a racket* in token 7 was heard as 'a lot kit' even though the speaker pronounced [ræ] at the beginning of the word. In each of these examples, word stress seems to be the main innovative feature. Tokens 4 and 6 were produced during conversations about the street scene. Because the speaker in token 4 pronounced the word *umbrella* quickly and with stress on the last syllable, the listener was not able to propose what had been said and instead replied that it sounded like a foreign language to him. Token 6 is a little different in that the non-standard grammar of *taking some photos **on*** something or the use of the article *a* instead of *an* before *orchestra* could possibly have contributed to the misunderstanding of *orchestra* being heard as 'a pesto'. Nevertheless, word stress was also implicated in the misunderstanding.

No.	Speaker	Listener	Word(s)	Heard a	Context
3	MCb1	FId2	circled	called	number thirteen (.) is **cirCLED**
4	FTh	MCb2	obstacle	??	use the fruit cart as the uh **obSTAcle**
5	FCb	MLa	umbrella	??	there is an **umbreLLA**
6	MCb3	FLa1	orchestra	a pesto	taking some photos (.) on a **orCHEstra**
7	MCb3	FLa2	a racket	a lot kit	there is **a raCKET**

Table 2. Instances of innovative word stress leading to misunderstandings
(Lewis & Deterding, 2018, p.170)

However, there were also instances in which non-standard word stress did not adversely affect intelligibility. Table 3 demonstrates a few of these cases. The *Said as* column shows the variant stress pattern that was produced. In each of these cases, listeners heard the word that the speakers intended. While many of the words such as *injured* and *mechanic* were only occasionally pronounced with unexpected stress, words like *calendar*, *guitar,* and *photographer* were often pronounced innovatively among the participants of this corpus with no impact on intelligibility. This supports Deterding and Kirkpatrick's (2006) observation that innovative stress patterns such as these are common among ASEAN speakers and do not seem to negatively impact intelligibility.

No.	Speaker	Listener	Said as	Context
8	FTh1	FIn2	calenDAR	on the calenDAR
9	FMm	MLa	GUItar	a GUItar (.) a GUItar
10	FMm	MCb2	photoGRApher	one man is photoGRApher
11	FIn3	MTh1	PHOtographer	there is a PHOtographer
12	MVn	FPh	inJURED	people (.) was inJURED
13	FTh2	MCb3	MEchanic	he looks like uh (.) maybe the MEchanic

Table 3. Instances of innovative word stress with no impact on intelligibility
(Lewis & Deterding, 2018, p.171)

From a preliminary analysis of the data, it is clear that ASEAN speakers sometimes utilize non-standard word stress patterns. While those innovations may be understood at times, there are other instances in which they hinder intelligibility. As the corpus of misunderstandings is analyzed further, it is possible that some predictability will be discovered as to which types of innovative stress affect intelligibility most in ASEAN conversations. Earlier research on the

intelligibility of word stress suggests that stress shifts that move leftward are less detrimental to intelligibility than rightward shifts (Field, 2005), that vowel changes significantly contribute to word stress misunderstandings (Cutler, 2015), and that a *word stress error gravity hierarchy* combining these two factors may estimate which word stress innovations will impede intelligibility more than others (Richards, 2016, p. 6).

Conclusion

As pronunciation is one of the most crucial features to evaluate when considering the intelligibility of ELF exchanges, and learner avoidance may influence language choices in naturally occurring settings, a discuss-the-differences task is an effective way of evaluating focused intelligibility. The combination of ELF, pronunciation, and intelligibility research methods in this type of research task produces corpora in which speakers are also listeners of the language, minimal background knowledge is needed, specific pronunciation features can be focused on, and language is contextualized for researchers.

Discuss-the-differences tasks are not without limitations, though, as the weaknesses in the drawings may not be discovered during the piloting stage, some researcher bias could remain, participant feedback may not always be reliable, and too much context might be given. In the drawings used in this current research, it was not until the research had been completed that the researcher realized that some of the objects in the pictures were misunderstood. Some objects such as the *vacuum cleaner* were not familiar to the speakers, while other items were misinterpreted such as the *fruit seller* regularly being mistaken for a 'baker'. After listening to the recordings, the researcher also realized small changes that would have made the research stronger. Changing the fruit cart to a vegetable cart would have encouraged the production of more polysyllabic words, and placing fewer items in the pictures would have made the task less daunting and narrowed down the vocabulary. Lastly, the researcher is not sure that allowing them to see their partner's pictures before obtaining their feedback was ideal. Though it created motivation for the speakers, those participants with sharp memories could have possibly remembered their partners' pictures and based their feedback on what they remembered.

As both listeners and speakers will rarely fully remember their conversations, their feedback may not always be reliable. Not only will the speaker sometimes believe they said something that they did not say but they will also occasionally not understand their own speech. Similarly, listeners may not fully recall what they heard at the moment of recording and their feedback might be based mostly on what they hear at the time of the feedback. So, on the day of the recording, they may have heard one thing, yet they might reply with a different answer on the feedback survey.

The researcher also acknowledges that some bias on her part still remains as she was the one to determine which clips were sent out for feedback and she will be the one to determine which feedback is reliable. Though the researcher attempted to base her follow-up requests on instances of listener clues (silence, laughter, or changing of topics), it is still possible that her perspective of what determined unexpected pronunciation influenced her selection of which clips listeners evaluated. Also, she will be the one to decide if misunderstandings occur. For example, a Cambodian speaker observing the organized street scene stated, as in (3).

(3) *1* MCb4 : and i saw a man (0.4) like a **drunk** man (.) a poor man (0.6)
     2   that uh drinking **alcohol** (0.5) maybe (0.6) in in uh near near
     3   the electronic shop (0.4) uh (.) and an old man (.) maybe
     4   is a (0.8) **begGAR**

The listener from Myanmar, who had informed the researcher that her internet did not function well, replied that she heard 'young' instead of *drunk* in line 1, 'coffee' instead of *alcohol* in line 2, and 'girl' instead of *beggar* in line 4. The researcher can understand how

drunk could be heard as 'young' and *beggar* with stress on the second syllable as 'girl'. However, she is not certain that *alcohol* was really heard as 'coffee'. The listener probably could not see the bottle in her chaotic picture and may have just guessed a drink from context. The researcher will be the one to determine if all of the above differences should be considered as misunderstandings.

Finally, while it is important to try to reduce the amount of background knowledge needed for a conversation in intelligibility research, Thomson (2018) warns that using pictures in intelligibility research might result in "giving a false sense of intelligibility" (p. 14). Researchers using discuss-the-differences tasks should be aware that some utterances may be understood primarily based on what they see (using top-down processing strategies) and not on what they hear (using bottom-up, word-level processing). Thus, if possible, researchers should try to use minimal pairs or near minimal pairs in their pictures which would encourage more focus on listening. For instance, if researchers want to study consonant clusters, they might consider having picture A with a *snail* on the ground and picture B with a *nail* on the ground, or picture A with a *clown* on the *floor* and picture B with a *cow* on the *four* (instead of the *three* or *six*). True word stress minimal pairs such as *INsight* and *inCITE* are limited (Cutler, 2015) and often more advanced lexically, so they were difficult to include in the current research. However, numbers such as *14* and *40* were included with this in mind as word stress is often taught as a way to distinguish between them.

While pictures have been used in some research studies on intelligibility and pronunciation in the past, discuss-the-differences tasks have not specifically been implemented until now. More research needs to be done on how well these types of tasks help researchers analyze other types of pronunciation features. The results of discuss-the-differences tasks also need to be evaluated in conjunction with other types of research, such as naturally occurring data and classroom observations.

References

Breiteneder, A., Pitzl, M., Majewski, S., & Klimpfinger, T. (2006). VOICE recording methodological challenges in the compilation of a corpus of spoken ELF. *Nordic Journal of English Studies*, *5*(2), 161–187.

Cruttenden, A. (2014). *Gimson's pronunciation of English* (8th ed.). New York, NY: Routledge.

Crystal, D. (2012). A global language. In P. Seargeant & J. Swann (Eds.), *English in the world: history, diversity, change* (pp. 152–177). New York, NY: Routledge.

Cutler, A. (2015). Lexical stress in English pronunciation. In M. Reed & J. M. Levis (Eds.), *The Handbook of English Pronunciation* (pp. 106–124). West Sussex: Wiley Blackwell.

Derwing, T., & Munro, M. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly, 39*, 379–397.

Deterding, D. (2013). *Misunderstandings in English as a lingua franca: An analysis of ELF interactions in Southeast Asia*. Berlin: De Gruyter Mouton.

Deterding, D., & Kirkpatrick, A. (2006). Intelligibility and an emerging ASEAN English lingua franca. *World Englishes, 25*(3), 391–410.

Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly*, *39*(3), 399–424.

Folse, K. (2006). *The art of teaching speaking: Research and pedagogy for the ESL/EFL classroom.* Ann Arbor, MI: University of Michigan Press.

Hancock, M. (1995). *Pronunciation games.* Cambridge: Cambridge University Press.

Jenkins, J. (2000). *The Phonology of English as an International Language.* Oxford, UK: Oxford University Press.

Jenkins, J. (2007). *English as a lingua franca: Attitude and identity*. Oxford, UK: Oxford University Press.

Jenkins, J. (2015). *Global Englishes* (3rd ed.). New York, NY: Routledge.

Kang, O., Thomson, R., & Moran, M. (2018). Empirical approaches to measuring intelligibility of different varieties of English in predicting listener comprehension. *Language Learning*, *68*(1), 115–146.

Kirkpatrick, A. (2010). *English as a lingua franca in ASEAN*. Hong Kong: University of Hong Kong Press.

Lewis, C., & Deterding, D. (2018). Word stress and pronunciation teaching in English as a lingua franca contexts. *CATESOL Journal*, *30*(1), 161–176.

Levis, J. M. (2005). Changing context and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, *39*(3), 535–543.

Long, M. (1989). Task, group, and task-group interactions. *University of Hawaii Working Papers in ESL, 8*, 1–26.

Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility and intelligibility in the speech of second language learners. *Language Learning*, *49*(1), 285–310.

Munro, M. J., & Derwing, T. M. (2015). Intelligibility in research and practice: Teaching priorities. In M. Reed & J. Levis (Eds.), *The handbook of English pronunciation* (pp. 375–396). Malden, MA: Wiley Blackwell.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985). Speaking clearly for the hard of hearing: Intelligibility differences between clear and conversational speech. *Journal of Speech & Hearing Research*, *28*, 96–103.

Rajadurai, J. (2007). Intelligibility studies: a consideration of empirical and ideological issues. *World Englishes, 26*(1), 87–98.

Richards, M. G. (2016). *Not all word stress error are created equal: Validating an English word stress hierarchy.* (Unpublished doctoral dissertation). Iowa State University, Ames, Iowa.

Seidlhofer, B. (2001). Closing the conceptual gap: The case for a description of English as a Lingua Franca. *International Journal of Applied Linguistics*, *11*, 133–158.

Seidlhofer, B. (2011). Understanding English as a Lingua Franca. Oxford: Oxford University Press.

SoGoSurvey. (2017). SogoSurvey [Online survey tool]. https://www.sogosurvey.com

Thomson, R. I. (2018). Measurement of accentedness, intelligibility and comprehensibility. In O. Kang, & A. Ginther, (Eds.), *Assessment in second language pronunciation* (pp. 11–29). New York, NY: Routledge.

Thomson, R. I., & Derwing, T. M. (2015). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics, 36*(3), 326–344.

VOICE. (2013). *The Vienna-Oxford International Corpus of English* (version 2.0 online). https://www.univie.ac.at/voice/