Active Learning: Writing and Speaking about Listening in a Business English Context

Alan Simpson

Abstract

Video conference needs analysis was conducted between Japanese and Chinese contact lens researchers. This revealed that the Japanese researchers struggled to understand the Chinese English. Furthermore, they often did not articulate this lack of comprehension due to differences in position and experience. Therefore, the Japanese researchers, who were also English students, wanted to build their listening skills. So they developed action research cycles to audio record and transcribe parts of their video conferences, highlighting which phonological features were problematic. This pilot study helps our understanding of phonological pronunciation and perception within contextually constrained Business English as a lingua franca environments.

Keywords

Chinese English pronunciation; L1 influences; listening; intelligibility; Business English as a lingua franca; action research

Introduction

Business English as a lingua franca (BELF) started with Louhiala et al.'s (2005) seminal work comparing the English used by two Nordic companies in a corporate merger. It was

defined as how cultural backgrounds and first language discourse practices are revealed in business interaction. Other ground breaking work was Jenkins (2000) study of the pronunciation forms of English as a lingua franca (ELF) trying to identify a Lingua Franca Core of pronunciation features. Similarly, Seidlhofer's (2001) Vienna-Oxford International Corpus of English (VOICE) aimed to identify discrete lexical features of ELF. However, ELF has had its critics (Sowden 2012, O'Regan 2014, Dauer 2005), who argued against the implausibility of a standardized lingua franca system. Since then, ELF research has moved away from the focus on lists (Seidlhofer 2008), and evolved to focus more on the function of the exchange and the fluidity when negotiating meaning (Seidlhofer 2009). This means that there are hybrid cultural practices (Baker 2015), and ELF is beginning to focus more on the multilingualism and multi-competence of the community, rather than just the English language or Anglophone cultural norms (Jenkins 2015).

When Ehrenreich (2016) more recently described the key issues in BELF, she described the confidentiality that often limits the amount of research data available. Also, that ethnographic and multi-method research approaches give a more accurate qualitative description of the BELF interaction. In BELF contexts, power and social distance affect who speaks when and the topic of conversation. There is still the variation in phonology common in other ELF contexts, but also specific lexico-grammar, discourse and pragmatic styles which are appropriate to that industry or community-of-practice (Ehrenreich 2016). Relational talk is also an integral part of BELF communication and is thought of as more difficult than business or specialized technical talk (Ehrenreich 2016). Cooperation and collaboration make BELF exchanges successful. So participants who have more established

relationships develop sensitivity towards culturally specific conversational norms, more often than transient and ad hoc ELF encounters (Cogo 2015, Jenkins 2015).

This article starts with a needs analysis of a corporate video conference where BELF was used, including the type of company, interaction, biographical information about the participants, and the data collection methods, following Serafani et al.'s (2015) needs analysis methodological recommendations. This revealed that there was a need to develop a listening based curriculum, before pragmatic moves can be considered. Therefore, the main focus is on the phonological factors which caused comprehension difficulties, and how the students investigated these issues. Finally, some implications for future Business English as a (multi) lingua franca research and training will be described.

Needs analysis

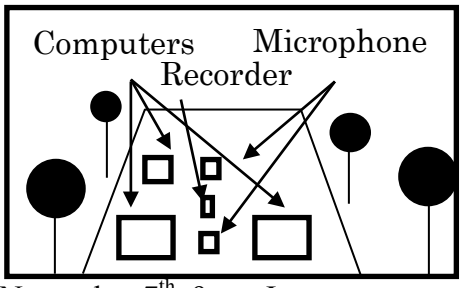
Initially a video conference was observed and analysed to identify communicative needs. This was followed by open and inductive interviews and follow-up emails with the Japanese participants. The project was conducted in a Japanese research facility of an American contact lens company. Monthly video conference calls were made between the head office in California and the research facility in Japan. The video conference was chosen as the main source of the Japanese students' needs analysis because of convenience and access. The students had just finished a one year English training contract so the author had a prior relationship, and when deciding what kind of future curriculum they would like, asked if it would be possible to observe their monthly video conference meeting. After some negotiation and understanding of the potential benefits, approval at the local level and with their Chinese manager in the U.S. was given, and confidentiality agreements were

made. Therefore, all references to names are pseudonyms, technical information has been removed, and approval to publish this paper has been given.

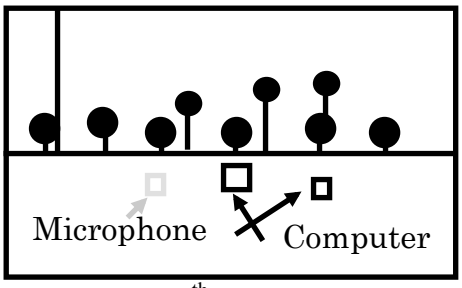
The video conference schedule and agenda was arranged by the chairperson in the U.S. There were nine participants on-screen and several off-screen in California, as shown in figure 1, four were Chinese. The Chinese chairperson, Ying as she'll be referred to, created an agenda, opened and closed the meeting, guided from one speaker to the next, and managed any discussions. Ying was the senior manager of chemistry and responsible for contact lens material development and the team of researchers in the U.S. She graduated from a Chinese university then studied her PhD in Japan, before moving to Canada for 5 years, and then the U.S. for 9 years. Bo, another active participant in the discussions, was a senior researcher working in the contact lens formulations group. He was from North Eastern China, but his language background, and the amount of time he had spent in the U.S. was unclear. However, he did admit, when asked at a later date by a Japanese researcher, that his English was not as strong as other researchers in California, and suggested that the Japanese students should study more proficient speakers.

On the Japanese side, there were three contact lens researchers, and the author, also shown in figure 1. Jiro was a semi-retired director, Hisa, a manager and chemical scientist, and Hitomi, a contact lens researcher. There was a small video screen in the top right of the video conference screen showing the participants, and the main screen showed the presentation slides. Both experienced and inexperienced researchers in Japan and California gave presentations on research progress, and the discussion focused on contact lens material development approaches.

Figure 1: The video conference screens



November 7th, 9am: Japan



November 6th, 6pm: California

Some of the initial difficulties for the Japanese staff included: bad sound quality due to the use of only one microphone in California; receiving the presentations only minutes before the video conference started; ambiguity caused by some presenters being off-screen in the U.S.

Investigating specialist discourse

The following sequential analysis is based on Conversation Analysis principles, adapted to show phonemic transcriptions of problematic phrases. This isa similar approach to O'Neal (2014) and Matsumoto (2011) who used phonemic Conversation Analysis transcriptions to show how interlocutors negotiated meaning when pronunciation features caused a breakdown in communication. The transcription conventions are shown in the appendix.

In the following two extracts, Hisa is the senior researcher in Japan and of a similar status to Bo. The discussions are about monomer synthesis, however, the response tokens were not a clear indication of comprehension.

Extract 1: Continuers

48. Hisa: What happened?
49. Bo: (3.0) Uh, the fact is that, the fact that the lens de-
50. degraded, and eh, eh some, some, eh, some type (xxx),

51. and eh, and also lost us our weight./lɔsɜsæʊweɪ/
52. You know?
53. Hisa: Hmm

The ‘Hmm’ response was inconclusive as to whether there was comprehension at this point, or it was just used as a continuer. However, contextual comprehension was shown later in the same sequence by Hisa because he asked a question.

Extract 2: Showing comprehension

63. Hisa: Do you remember I have already synthesized the macromer
64. Bo: Yes, uh, (we want to get some macromer from you so
65. that) you said so the formation of the (xxx), eh so,
66. (perhap), maybe you can ship /ʃep/ me some (xxx)?
67. (7.0)
68. Ying: Do you have the material available?
69. Hisa: Ah, ((laughs)) maybe, uhm, less than 3 grams, now.
70. Maybe. I will check it.

In line 67, the seven second delay is interactionally relevant. Bo has asked Hisa a question. However, Hisa does not respond. Therefore Ying orients to the silence as an indication that Hisa does not understand because she reformulates the question to help Hisa to make an answer. However, Conversation Analysis does not clearly prove whether Bo’s pronunciation was problematic for Hisa. Similarly it does not reveal whether the ‘hmm’ response token in extract 1 was to acknowledge comprehension or to ‘let it pass’ (Firth 1996) and try to guess the meaning from context later. Therefore, follow-up active

interviews and follow-up emails were used to co-construct, verify, and develop a deeper ethnographical understanding (Mann 2011, Basturkmen 2010).

Hisa said, “Bo is more difficult to understand than other people. Especially his voice production is not so clear.” This was supported by Hitomi and Jiro who said that, they could not understand the Chinese pronunciation and felt that some speakers spoke too quickly. In addition, Hitomi struggled with some of the technical vocabulary and said that some of the required actions were not clear, so follow-up emails were needed to clarify action points. All of the students said that listening was the most difficult issue. Even with the support of presentation slides, there was a lack of comprehension especially when listening to Ying and Bo.

This was in contrast to Louhiala et al.’s (2005) findings from their Nordic business context that phonemic and syntactic levels of communication rarely hindered communication. In this context communication continued but with impeded comprehension. Ehrenreich (2016) acknowledged that phonological variation, power issues in a business context, and language ability can affect the pragmatics of clarification. In this context, the lack of clarification and negotiation of meaning was revealing.

The students and aims

The class participants, as shown in figure 2, were the three needs analysis video conference participants, Hiro, Jiro and Hitomi, two younger contact lens researchers, Maki and Kaori, who participated in subsequent video conferences, and a new employee, Kei, who had recently graduated from university. They had a range of English abilities. They all had

chemical engineering backgrounds, and a familiarity with the technical vocabulary. However, Hisa had a doctorate and twenty years of experience, and Jiro had a full career of experience of the vocabulary, processes, and background in this field. On the other hand, Kei, who had just joined the company and was learning about the job, was not able to join the video conferences yet. However, he was encouraged to learn about these processes by studying the speaker's presentations and transcribing the audio recordings.

Figure 2: The students

After listening to some samples from the needs analysis video conference, the students agreed that the recordings were noisy, difficult to connect key words and understand the gist of what was being said. Therefore, the following strategies were agreed upon:

- Listen to authentic recordings of the video conferences, so that familiarity with pronunciation features and listening comprehension could be improved, especially to Ying and Bo

- Transcribe and discuss extracts
- Record Californian staff introduction interviews to study their basic pronunciation without the added complication of technical vocabulary or ambiguity because of the video conference system

The aim of this research was to discover which Chinese English pronunciation features were difficult to recognise when the Japanese students listened to them. Deterding (2013), in his study of misunderstandings in Asian ELF speech, also included instances when there was no apparent breakdown and the conversation seemed to progress smoothly, but one of the interactants subsequently reported that there were some things they had not understood. Giving minimal backchannels such as *mm* or *yeah*, are common ways of using the 'let it pass' strategy (Firth 1996). The primary role of such backchannels is to confirm that the listener is following what the speaker is saying, but they can also be used to conceal that the listener does not understand and is trying to let the conversation flow (Deterding 2013). This is a completely legitimate conversation strategy, and is common among lower level proficiency speakers (O'Neal, personal communication). There are not always obvious signs within the interaction of comprehension difficulties, so if there are ambiguous discourse markers, to signify comprehension, then ethnographical approaches must be taken to discover the listener's understanding.

The Listening Process

English speakers in inner circle countries (Kachru 1985, 1992, 2005) are not necessarily more intelligible than outer or expanding country speakers (Smith and Rafiqzad 1979;

Smith and Bisazza 1982; House 2003; Shaw et al. 2009), because they do not pronounce every sound (Cogo 2009; Deterding 2013), and it is difficult to define a standard native pronunciation. Therefore, when non-native English speakers are communicating with other non-native English speakers, the relation to native speakers becomes irrelevant, and mutual intelligibility becomes the crucial interactional factor. Intelligibility is the ability to understand the words articulated through a pronunciation (Munro, Derwing & Morton, 2006; Nelson 2011), and mutual intelligibility is a judgement made in relation to both the speaker and listener. It is a dynamic, negotiated process, embedded in context, and is affected by the listener's familiarity and attitude to the accent they hear as much as the speaker's pronunciation (Goh 2000; Rajadurai 2007; Rogerson-Revell 2007). Non-native contexts also often involve a lack of common cultural knowledge, so there is more focus on meaning, rather than grammatical correctness. Zhu (2015) when discussing the negotiation of meaning in intercultural contexts, highlighted that divergence from the other speaker could be shown by not adjusting accent, speech rate, pausing or turn taking patterns, utterance length, and gestures, and not taking into account the listeners' perceived linguistic ability, interpersonal needs or goals of communication. This could also lead to a lack of confirmation checks, self-repair, and interactive repair (Mauranen 2006). How pronunciation features are recognised and interpreted are dependent on the experience and ability of the listener. These differences, combined with the sheer growth of non-native English speakers around the world creates a strong need for more research into local issues especially in BELF contexts (Bargiela-Chiappini & Zhang 2013), and creates a demand for new pedagogical approaches.

When people listen to a spoken sentence, they think that they are hearing sequences of discrete sounds and words that correspond to vocabulary in their memory. However, this is an illusion created in the mind, because when listening to an unfamiliar foreign language, a continuous stream of speech is heard without knowing where the words begin and end. The listener is not sure how to segment linguistic units into words and so relies on lower level or bottom-up processing (Collins & Mees 2013). There are cognitive constraints on working memory. Echoic memories are stored for about four seconds in the auditory cortex (Darwin et al. 1972). Therefore, if there is only one opportunity to listen to live interaction, other strategies, pragmatic or higher level processing is required (Richards 1983). What listeners select for processing becomes crucial. They need to focus on semantic clues and chunk larger units into meaning which can be dealt with quicker, and must resist the temptation to translate (Vandergrift 1998a). Another factor is the student's ability or willingness to keep on listening and not be distracted by unfamiliar words (Goh 1998).

There are two contrasting views of speech perception. The first is the bottom-up approach, where the listener builds up a representation of what was said from the lowest phonemic level, for example, /t//ɛ//s//t/ into longer phrases or chunks, from which we can derive meaning (Anderson & Lynch 1988).

The second is to build a top-down perception by using prior knowledge and expertise to guess, predict and fill in the meaning (Pinker 1994). Hearing research suggests that both higher-level interpretations and lower-level input representations must be simultaneously available in order to support successful speech perception and learning

(Davis and Johnsrude 2007), and this is supported by psycholinguistics (Anderson and Lynch 1988; Flowerdew 1996; Lynch 1996). Having said that, many even high level non-native speakers prefer to use bottom-up strategies rather than contextual ones (Jenkins 2000). It seems like a default fallback position. This puts a higher reliance on the acoustic signal, trying to match it with some known vocabulary items from memory. However, their guess then can be quite different from the sound. They are processing after listening rather than anticipating beforehand. "Learners are less sure of the discourse, syntax, and typical vocabulary used in the discussion of a particular topic", (Brown 1990: 60). When something does not make sense, they are unable to decide whether the speaker used an unfamiliar word, they have misheard, or the speaker has unfamiliar pronunciation (Conrad 1983; Jenkins 2000; Pickering 2006; Deterding 2013), because of a lack of contextual inferencing skills.

L1 phonological transfer

Some breakdowns in ELF communication are because of L1 phonological transfer to the L2 (Jenkins 2000; Simpson 2014). There are important differences between the phoneme sound systems of different languages. For instance, Japanese people at first hear no difference between the contrasting English phonemes /r/ and /l/, because they do not exist in Japanese. It is difficult to perceive phonemes when there is no equivalent in the learner's L1 (Guion et al. 2000). Therefore, "articulation must be learned before auditory differentiation can be realized" (Jenkins 2000: 34). Pedagogically, it is crucial to accept L1 phonological transfer as a universal, then identify which areas are affecting intelligibility, and address those " (Jenkins 2000: 104).

In Chinese, there is no differentiation between long and short vowels. Therefore, Chinese speakers of English often pronounce short vowels as long vowels (Qian 2011). Chinese morphemes are usually made of a consonant plus a vowel, with no consonant clusters, and words usually end in a vowel, the same as Japanese. Chinese is a syllable timed language, while English is a stress timed language. For Chinese rhythm, the production of every syllable takes virtually the same amount of time, while in English the stressed syllable takes longer to produce. Therefore, some Chinese speakers over-articulate unstressed English syllables and words, or under-articulate consonant clusters depending on their language experience and how carefully they are speaking. If key words, that contain new or important information, are not stressed or pronounced more slowly and clearly than other words in English, it makes hearing them more difficult (Zhang & Yin 2009).

The Lingua Franca Core

The Lingua Franca Core (LFC) is a set of pronunciation features which were identified by Jenkins (2000) as being necessary for achieving mutual intelligibility between non-native English speakers. There have been scholars who have supported this simplified pronunciation model, such as Gilbert and Levis (2001), Hewings (2001), Pow (2002), Grazzi (2003) and Walker (2001), who showed how the LFC can be applied to make pronunciation teaching simpler. However, Sobkowiak (2008) suggested that the LFC promotes errors, which Jenkins (2008) refuted because that is in relation to a native speaker pronunciation standard, which the LFC is not.

It is important to appreciate that the LFC features are not a pronunciation standard, such as General American (GA) or Received Pronunciation (RP), but as a guide for

understanding how and why certain features cannot be understood.

According to Jenkins (2000), the most important areas in the LFC to preserve mutual intelligibility are:

1. Most consonant sounds

/r/ is rhotic (or pronounced), like GA, /t/ is closer to RP (rather than the flapped /t/-/d/ in GA, and /p/, /t/ and /k/ should be aspirated, with a burst of air. The two consonant phonemes, /θ/ and /ð/, are not included because for Jenkins, they did not cause any phonological confusion.

2. Appropriate consonant cluster simplification

In a medial cluster there is often elision such as in lis<u>t</u>en and pos<u>t</u>pone. In the initial cluster there should never be elision, for example a Taiwanese speaker's pronunciation of the word *product* is unintelligible as [ˈpɑdʌk]. Final consonant clusters often have elision: scrip<u>t</u>s, fac<u>t</u>s, secon<u>d</u> class. So in the LFC the initial consonant cluster is seen as crucial.

3. Vowel length distinctions

Vowel length is reasonably stable across English varieties, whereas vowel quality is not. Therefore, vowel length rather than quality is more important for intelligibility (Jenkins 2000). The important part of Jenkins argument is that many L1 speakers have different vowel qualities, so it is not necessary for learners to use their vowel qualities as reference models, whereas vowel length has distinctive patterns, and if the speaker deviates from these patterns, there is a higher risk to intelligibility. Diphthongs also differ widely in quality among native speakers and so they too cannot be given high priority in L2 teaching.

Therefore, the LFC omits monophthong and diphthong quality as core items, because they are unteachable in a classroom context, as long as they have consistent length.

4. Nuclear stress

A lot of pronunciation pedagogy is based around the importance of suprasegmental variation (i.e. prosodic, stress, rhythm, and intonation). There are two kinds of stress: word stress and nuclear stress. Word stress is the stronger syllable in a word, whereas the nuclear stress is on the most prominent syllable in an intonation group (Collins and Mees 2013). The nuclear stress is used to identify the most important information. For non-native English speakers, the greatest phonological obstacles to mutual intelligibility according to Jenkins (2000) are deviant LFC sounds in combination with misproduced nuclear stress. Outside these LFC features, Jenkins said that L2 variation should be regarded as regional accent variation, and this would allow much greater freedom than GA, RP or some L2 varieties for individuals to express their identity and to accommodate their receivers. As Jenkins said, (2007), these are not fixed features and are open to further research based evidence.

Methodology

The pilot study methodology was based on student led action research cycles, developed from Zuber-Skerritt (2001), which revolve around the principles of developing a plan of action, then acting, observing and reflecting. Over a period of six months, a total of five video conferences were recorded. The first one was the needs analysis observation. This was transcribed by the author, then he selected twelve extracts of potential interest to the students, including technical explanations, feedback, and interaction, including instances

which seemed hard to understand. Then a further four video conferences were audio recorded by the students, and they selected short two minute extracts, based on their interests and needs. The audio files were then shared with the other students and everyone transcribed them as homework. In addition, two California based members of staff, Bo and Ying were interviewed and recorded by a student while she visited the U.S. on a business trip. Bo and Ying were chosen because they were the prominent members of staff, and their Chinese English was difficult to understand. They took up a lot of the video conference interaction time, by clarifying research methodology, advising, giving progress updates and highlighting action points. The interviewer, Kaori, asked some biographical information specifically about job descriptions. Then she asked the interviewees to read her technical presentation slide script. So that the students could familiarize themselves with the speakers' pronunciation patterns.

Many dictation based methods of writing transcriptions to improve listening skills are based on audio files which are not intended for the listeners (O'Neal 2014). However, this was not the case in this context. The audio files were not just intended for the listeners, but were also consequential for the actions required in their jobs. This methodology is similar to Deterding (2013), who analyzed data with the help of self-reported transcripts. The students were also introduced to the Conversation Analysis concept of a data session, which is a meeting of data researchers in order to discuss some recordings and transcripts (ten Have 1999).

In class, a different student co-chaired each data session and all students listened again to compare their transcripts. The author acted as co-chairperson and was an active

participant observer on the transcription and data session processes. He compared the students' transcripts and compiled a corpus of problematic features. These problematic words or phrases were chosen because they were unintelligible for three or more students. For reference and to triangulate the results, the author's perception was also transcribed using *TypeIt,* http://ipa.typeit.org/full/, which is an online website used to type IPA phonetic symbols. The author's transcriptions of the speaker's pronunciation may include inaccuracies because they are simply based on his perceptions. The author's perception was important in trying to assist in deciphering what was actually said. American dictionary pronunciation was also used as a reference, because the speakers worked and lived in California. This was transcribed using *Lingorado,* http://lingorado.com/ipa/. The American phonemic spellings may have inaccuracies depending on the standard used by http://lingorado.com/ipa/. This American dictionary pronunciation was useful to understand how the Chinese English pronunciations vary, not as a reference for what is correct, but to be able to understand that variety so that it could be further compared to Jenkins (2000) lingua franca core. If there had been negotiation between the speakers to clarify what was said, then that could have been used as a data source (Matsumoto 2011, O'Neal 2014). However, lacking a negotiation of meaning and segmental repair, and as a listening training tool, then these references are required. After the research was completed, semi-structured individual interviews were conducted to discover their perspectives of the transcription and learning cycles (Mann 2011).

Analysis

Problematic Pronunciation

Bo changed his vowel sounds in *senior, macromer, caused and solve*. He also dropped his *r* in *for*mulation and *proper*ties, and added a *y* sound to the word *if*. His speech rate was measured at 100 wpm, which is below the 130wpm level highlighted by Rivers (1981) as slow. After talking to him face to face, Kaori said that she understood about fifty percent of what was said, which indicates that the speech rate was not the main problem.

The next six tables show examples of problematic consonants, vowels, assimilation, combinations, vocabulary, and unintelligible utterances. The General American pronunciation is a reference standard to what the researcher perceived was said. The problematic pronunciation features were identified as such by the students in their transcripts, and therefore clarified by the researcher to understand how the words were pronounced and why they were problematic for the students.

When Ying pronounced phrases containing /t/ it was problematic for the students, as shown in table 1.

General American pronunciation	Problematic Pronunciation	Description
What is needed /wʌt ɪz ˈnidəd/	What is needed [ˈwɑt̪ɪz ˈniːd↓ɪd]	The /t/ was pronounced as an alveolar stop [t̪], which sounded like a quick /d/ because it was between vowels. There was also a /dɪd / sound in needed, but the last part was very soft and difficult to catch for the students

And want do fast /ænd ˈwɑnt ˈduː ˈfæst /	And want do fast /ənd ˈwɔn ˈduː ˈfæst /	The /t/ was dropped because it was after a linking consonant and before another consonant. Also the lack of the preposition *to* was confusing.
Then it's really /ˈðen its ˈriːliː/	Then it's really [ˈðen ɪʔz ˈriːliː]	/t/ was replaced by a glottal stop because it was after a vowel and before /z/, so the students couldn't recognize what was said.
Sixty seventy /ˈsɪksti: ˈsevənti: /	Sixty seveny /ˈsɪksti: ˈsevən↓i: /	When a vowel comes after /nt/ the /t/ is dropped. The intonation also decreased at the end.

Table 1: Ying's /t/

The analysis of the pronunciation of *b, v, & w* revealed some interesting factors. Bo pronounced *vital, as* /ˈwaɪtə/, potentially because of the lack of the voiced labio-dental fricative /v/ in standard Mandarin Chinese. Furthermore, he dropped the /l/ because of a Chinese preference to finish with a vowel. However, surprisingly, Ying pronounced *switching as* /ˈsvɪtʃɪŋ/, clearly substituting /v/ for /w/. Further research, (http://answers.echineselearning.com/questions/2010-07/11/034121977EEEFWCDS.html retrieved April 3, 2015), revealed that pronouncing /w/ as /v/ is an L1 dialect influence which started in Northern China and Beijing and has become more popular throughout Mandarin Chinese, although it is not a standard feature of Chinese. Furthermore, Ying pronounced *obviously* as /ˈɔviːəsliː/, dropping the /b/. Initially, it was considered that this

was because there are no consonant clusters in Chinese, so Ying simplified the initial consonant cluster in the word. However, this can also be a feature of American English and could be because Ying has lots of experience of American English (O'Neal, personal communication).

Table 2 shows some of Bo's vowels which caused confusion, including shortening and lengthening vowels, vowel variation, and epenthesis or adding an extra sound between two sounds.

General American pronunciation	**Problematic pronunciation**	**Description**
very /ˈvɛriː/	Veri well /veri/ (or vewi)	Vowel shortening, /r/ was pronounced but the vowel length was shortened, making it sound like *vewy well*.
not care /ˈnɑt ˈkɛr /	nawt care /næʊt ˈker /	Vowel lengthening, /ɑ/ changed to /æʊ/
country / ˈkʌntriː /	cowntree /ˈkæʊntriː/	/ʌ/changed to/æʊ/
but anyway / ˈbʌt ˈɛniːˌweɪ/	but ainyway /ˈaɪniweɪ/	/ɛ/changed to the fronting /aɪ/ diphthong
senior / ˈsiːnjər /	seneor /ˈsiːniːɒr/	/jə/->/iːɒ/
stability /stəˈbɪlɪtiː /	stibalidy /ˈstɪbɑːlɪdiː/	Vowel variations, /ə/->/ɪ/, /ɪ/->/ɑ/ Central vowel moved forward and front vowel moved back
permeability /ˌpɜrmiːəˈbɪlətiː/	permeabality /pɜːmiːæˈbɑːlɪtiː/	/ɪ/->/ɑ/ Central vowel moved back
skills /ˈskɪlz /	skalls /ˈskɑːlls/	
is very clear	..your presentation yis vewy	Epenthesis, /ɪ/->/jɪ/

/ ɪz ˈvɛriː ˈklɪr /	clear /jɪs vewi/	
issue /ˈɪʃuː /	This is a yissyu about /əˈjɪːʃjuː/	wrong article use, and /ɪ/->/jɪ/
Eh if there are something wrong / ɛ ɪf ðeər ɑr ˈsʌmθɪŋ rɔŋ/	eh yif dere are someting wrong /ɛ ˈjɪf dɛr ɑr səmtɪŋ rɒŋ/	/j/ is used to link the vowels /ɛ/ & /ɪ/

Table 2: Bo's vowels

There were instances of vowel shortening, lengthening, inconsistent vowel variation, and epenthesis, or adding extra sounds. *Not* and *country* were pronounced using the backing and closing diphthong /æʊ/, not monophthongs. These changes in Bo's vowel pronunciations often involved moving the place of articulation. In the phrases *a issue,* /əˈjɪʃjuː/ and *Eh if* / ɛ ˈjɪf /, an extra /j/ was inserted as a linking sound to separate the vowels, called y-gliding**.** However, in the phrase *your presentation is* /jɪs/ *very clear,* the inserted /j/ was between two consonants. This pronunciation characteristic was common for Bo, although not all instances were problematic for the students. These examples suggest L1 cross-over because in Chinese, words do not start with vowels.

Some assimilation examples from Ying and Bo, which were difficult for the students to understand, included, /θ-ð/, blending vowels, and joining the same sounds, as shown in table 3.

General American pronunciation	Problematic pronunciation	Description
What's the issue? / wʌts ðiː ˈɪʃuː? /	Whaseyissue / Whasiː ˈjɪːʃjuː/	Lagging assimilation involving /s/ & /ð/,
contents there /k ɑnˈt ɛnts ˈðeər /	contentsere /kənˈtentser /	
Now we are doing /ˈnaʊ wiː ɑr ˈduːɪŋ/	Now we are doing /ˈnæʊwiɑr ˈduːɪŋ/	*Now we are* blended together.

Table 3: Problematic assimilation

It was difficult for the students to tell if Bo had made a mistake, used a new word, changed his pronunciation or joined his words together. This was often complicated by a combination of factors, as described in table 4.

General American pronunciation	Problematic pronunciation	Description
Now lately we have /ˈnaʊ ˈleɪtliː ˈwiː ˈhæv/	Now lately [ˈnæʊ ˈleɪʔliː]	There was a glottal stop after the vowel and before the consonant, and there was an /r-l/ Japanese student sound perception problem.
re-scope this /riˈskoʊp ðɪs/	re-scope this /riːskoðɪs/	This was said softly and the consonant cluster was shortened.
I also do	I also choo- do	Self-correction
read / ˈriːd /	reed /rˈiːd/	Unusual stress
Shelf life test	Shelf life test	New vocabulary

Table 4: Problematic combinations

There were issues on both the speakers' and listeners' sides. On the listeners' side the students were unfamiliar with glottal stops, the contrast between /l/ and /r/, self-corrections, and cut-offs.

The next listening sample contained important contact lens research feedback and was difficult to understand. Table 5 shows the students' and researcher's perception of what was said.

General American pronunciation	**Problematic Pronunciation**	**Description**
reaction [ɻiˈækʃən].	/riˈækʃən/	For two students, *reaction* sounded like *direction*. /r/ was pronounced in a central approximant manner in the post-alveolar region, realized as [ɻ]. /d/ was also articulated in the post-alveolar region, but pronounced as a plosive stop. Note: only understanding the end of the word for Kei was similar to Kaori six months previously, which would suggest that it was a stage of her listening skill development.
vinyl /ˈvaɪnəl/	/ˈwaɪnʊ/	/v/ was pronounced as /w/, and Chinese words end with vowels, which accounts for dropping the /l/.

influence too much /ˈɪnfluənst tu mʌʧ/	/ɪnfluəns tu matʃ/	There was equal stress on each syllable in *influence* and vowel variation in the word *much*. The difference between *so much* and *too much* had to be negotiated in the classroom discussion. It was finally agreed that in this context, a negative factor was being discussed so *too much* was the most appropriate phrase.
is acrylate / ɪz ækrileɪt/	/ɪz haɪprileɪt/ or [ɪz a^{h} ɪprileɪt]	The /a/ had aspiration and sounded like /h/. Both /k/ and /p/ had strong (fortis) voiceless articulation with plosive power. However, /k/ was pronounced with the back of the tongue (velar) and /p/ with the lips (bilabial). Two students and the researcher guessed it was *rate* rather than *late* at the end of the word. However, the more experienced students said that the topic was *is acrylate*.

Table 5: Bo's technical feedback

This cycle showed some of the complexities involved in the listening process. The first case related to the listener's sound perception and vocabulary, hearing *reaction* as *direction.* The second and third cases related to the speaker's L1 influences on his pronunciation of consonants, vowels, assimilation and word stress. In the third case, there was no stress, which usually does not cause problems in isolation (Jenkins 2000), however, combined with the pronunciation of *much* as /matʃ/ compounded comprehension difficulties.

Constructing Meaning

Comprehension improved when the students had a stronger vocabulary. Students who had more experience, Hiro and Jiro, understood technical words such as, *macromer, ether, reaction,* and *formulation.* The less experienced students were only using sound processing,

because of their lack of experience, vocabulary or prior knowledge. The word *acrylate* was specific vocabulary for this context, and involved a variation in the speaker's pronunciation and listener's perception. Extract 4 shows the discussion about this word and the sentence in which it occurred. Furthermore, it also reveals how the students tried to negotiate and construct meaning together.

Extract 4: Vocabulary and empowerment

1. Kaori: Let's move on to line G. Please sound.
2. Alan: Please play, ok ((plays audio))
3. Kaori: Hitomi, what did you think?
4. Hitomi: I heard the sentence, *the PEG*, I couldn't catch something,
5. *have similar structure*, *and the backbone is acrylate*.
6. Kaori: Ok thank-you. Ma-maki?
7. Maki: I heard the same *PEG*, dadada, *they have similar structure*,
8. *and their main backbone is acrylate.*
9. Kaori: So- (26.0) I- I ((doushio = what to do?)) May- hmm, many
10. person catched mm, *PEG* and *similar structure*, *backbone* and
11. *acrylate.*
12. Alan: Mmm
13. Kaori: On ((gestures the previous California slide)), ehm, they,
14. they, proposed acrylate type macromers, so he, he, he want to
15. say about acrylate type macromers, not hydrate or hyperate,
16. applirate.
17. Alan: (2.0) Yeah, I think some people wrote *h*, because sometimes
18. when Bo pronounces a word with a vowel at the start, like
19. /a/ in apple for example, then I think he'll put an extra
20. sound, because of the Chinese, so he'll say like *hacrylate*,
21. when he means *acrylate*. I think that is Chinese affecting

22. his English. I guess, but I think you're right, and the
23. word is *acrylate*. But that's probably why it's confusing
24. for me and Hitomi here.
25. Hitomi: ((nods)) I heard only sound so I transcript, I made the
26. transcript by listening. On here ((pointing)), on paper,
27. it's clear, acrylate.

Figure 3: Empowerment

This extract showed that Kaori led the conversation because she had a facilitator role and the author was no longer the gatekeeper of knowledge, because she had more content knowledge. She consolidated the participants' opinions and then gave a summary using the context and her experience.

However, the classroom discussions did not always clarify what was said, because some utterances were unintelligible, as shown in table 6. This table shows the researcher's perception as well as the student perceptions of what was said.

Author perception of what was said	Student perception of what was said	Author perception of General American phonetic transcription	Description
/ˈdevaɪd/	You tried to divide Tried to evaluate You tried imide You try to invite You try to why	You tried to divide /dɪˈvaɪd/	The stress was on the first syllable, and the vowel was more like /e/ than /ɪ/
/ˈkəntɑɪnɪdʒ/	Companion Continue Contains a Contains Containish Conpendage Can, can use a	Containing /kənˈteɪnɪŋ/	The stress was on the first syllable.
/ˈkʌlmɪneɪt/	Combinate carbonate conformat culminate confername we have made	Culminate /ˈkʌlmɪˌneɪt/	Low volume.

Table 6: Unintelligible utterances

In the first example, there was a problem understanding the vowel at the start, the nasal /m/ or /n/, the fricative /v/ or central approximant /w/, and the plosive /d/ or /t/. Also, there could have been elisions. Simply, there were too many unknowns for the meaning to be negotiated. It was unintelligible for the students. Similarly, in the second example, there was agreement that the word started with /kə/, but then there was uncertainty about whether

it was the nasal /m/ or /n/, the plosive /t/ or /p/, then some agreement about /æn/, /eɪn/, or /ɛn/, followed by a range of endings. *Continuing* was the most logical guess, even though that did not match the audio heard. Just as difficult was the third example, where people agreed that it started with /k/, then a vowel, then a range of guesses. This was also unintelligible. This audio sample was at a similar speech rate, 103 wpm, as other samples, but the level of vocabulary, assimilation, and variation was more difficult.

Discussion

Some scholars (O'Neal 2014; Louhiala et al. 2005), say that when there are breakdowns in lingua franca contexts, then pragmatic repair methods are more important than the reason for the breakdown. However, in this Asian BELF context, there were important factors which meant that the students did not actively engage in repair, including age, experience, and technical knowledge. Japanese culture has a tradition of being a high context culture, with the onus being on the listener (Hall E T. 1976, Gudykunst WB & Ting-Toomey S. 1988), with a high power distance factor (Hofstede G. 1983.) Therefore it is not a surprise to see these kinds of characteristics represented in the interaction through a lack of repair and passive listening techniques. Due to this, it is important to raise students' awareness of their own discourse practices, conventions and cultural preferences, and as intercultural experience increases then focus on correctness decreases and acceptance of differences improves (Wang & Jenkins 2016). Students and teachers focus more on language and form whereas professionals focus more on content (Ehrenreich 2009). For students to be able to understand more of the content, they should be taught how to be more assertive when they are comfortable doing so. However, at the same time, they also need to improve their

passive listening skills when they are not comfortable repairing these breakdowns in communication.

Therefore, the aim of this research was to use action research cycles to identify the pronunciation features which had an influence on student comprehension. At the end of the action research cycles, in one to one interviews, the students said that writing transcripts was a good way to build their listening skills, although they still wanted to build both their technical and informal vocabulary. All of the students talked about how they had learned some characteristics of Chinese English pronunciation, but understanding the video conference speakers was still very difficult. Jiro had a perception of English which was not being met, that he should practice listening to native speaker English to build up his listening skills before practicing listening to non-native speakers. This was similar to Bo's perception that the Japanese students should use a native pronunciation model. However Wang & Jenkins (2016) concluded that linguistic experience in intercultural communication between non-native speakers would improve comprehension of non-native accents. More extensive research is required to discover whether there is resistance towards 'imperfect' non-native models. From a pedagogical perspective, it is important for non-native speakers to pronounce carefully, and try to clarify and accommodate differences, rather than simply thinking that their pronunciation is not as clear as native pronunciation, whereas in many cases the opposite may be true.

During this course, the students said that they thought that part of the reason understanding Bo was difficult was due to his speaking rate. However, the problem was not his speed, because his speed was only around 100 wpm. Fast speech is above 220 words

per minute (wpm) and slow is below 130 wpm (Rivers 1981). The problem was being able to recognize when one word stopped and another started. Some students struggled to define the word boundaries and segment the data stream into words, and this created a false sense of speed. Furthermore, when a hesitation, false start, cut-off, mistake, or self-correction was produced, the students did not have the flexibility or experience to accept or accommodate these features, but they did raise their awareness. Kaori explained that it was difficult because different speakers had similar pronunciation, but some sounds were unique, like the way Bo pronounced /ɪ/ as /ɑ /. Jiro said that it depended on the person, some Chinese English speakers speak slowly and clearly and are easy to understand, but Ying and Bo were more difficult. The students also differed in how they approached the listening task. Hisa, Kaori and Hitomi preferred top-down inferencing strategies when they were listening, whereas Jiro, Mako, and Kei preferred bottom-up strategies. This led Jiro, in particular, who had good experience and knowledge of vocabulary, to hold onto unknown words in his short term memory for too long, and was not able to jump over unknown phrases and continue to follow the data stream. Listening strategies were variable between individual students, which is similar to Field's (2004) conclusions. Therefore, for the students who preferred bottom-up decoding, more inferencing and self-monitoring strategies would help (Renandya & Farrell 2011).

The phonological features which caused comprehension difficulties are shown below.

Problematic pronunciation features		
Non-rhotic /r/	/t/ variation, glottal stops, /t-d/ energy assimilation	/l-r/ speaker pronunciation and listener perception
/b v w/ variation	Consonant cluster simplification	Vowel quality
Vowel length	/j/ epenthesis	Assimilation, especially /θ-ð/
Nuclear stress	Hesitations, cut-offs and self-corrections	Combinations of these features

Table 7: Summary of problematic pronunciation (and communication) features

The students struggled to understand the pronunciation of Ying's *t's.* Jiro said that dropping or changing *t* and *r* were not difficult. However, the students' transcripts showed that they struggled to recognize these features. The difficulty for Kaori to differentiate between *l* and *r* was because of her Japanese mother tongue. The Japanese language possesses one liquid consonant, which is most often realized as an alveolar tap [ɾ]. English has two: an alveolar lateral approximant /l/ and rhotic consonant centred on [ɹ] (Hallé et al. 1999). English /l/ is perceived by Japanese listeners as more similar to Japanese [ɾ] than English [ɹ] (Guion et al. 2000), so *lately* was heard as *really* because the /l/ was heard as /r/ and there was a glottal stop instead of /t/. Hisa said, "it is difficult to catch words we cannot pronounce", which supports the case for integrated pronunciation and listening practice.

Bo often varied his vowel length pronunciation. Vowels are usually shortened before fortis consonants in native speaker varieties of English, such as *not*, / ˈnɑt /, and *caused* / ˈkɔzd /. However, Bo pronounced the monophthongs as diphthongs, such as *country* /ˈkæʊntriː/, *anyway* /ˈaɪniweɪ/, and *senior* /ˈsiːniːɒr/. Vowel variation included

stability changing to /ˈstɪ̲bɑ̲ːlɪdiː/, *permeability* changing to /pɜːmiːæ̲ˈbɑ̲ːlɪtiː/, *skills* changing to /ˈskɑ̲ːlls/*,* and *macromer* changing to /ˈmɪ̲kromər/. More frequently, the front vowel /ɪ/ moved backwards to /ɑ/. However, the exception was the central vowel /ə/ moving forward to /ɪ/, in the word *stability,* /ˈstɪ̲bɑːlɪdiː/. Therefore, his vowel quality was not easily predictable.

A combination of vowel and consonant variations were expected, yet being able to associate them with the Chinese pronunciation patterns was not. *Very* was a good example because Bo not only shortened his final vowel sound but also /r/ sounded like a compressed-lip velar approximant /w/. This confusion between /r/ and /w/ occurs widely in South China and Hong Kong (Deterding 2013), which shows that variations are not just an L1 influence but also depend on the educational, social and dialectal background of the speaker. This was also the case with Ying's substitution of /v/ in place of /w/, in the word *switching*, which contrasted with Bo's substitution of /w/ in place of /v/ in the word *vital.* Similarly, Ying omitted /b/ in /ˈɔviːəsliː/. This was initially considered an L1 influence. However, it was discovered that this pronunciation is also common in the U.S. This showed that it was not simply an L1 influence or a lack of knowledge about English elision rules, but more generally due to Ying's education and experience. Therefore, listeners may struggle when listening to both native and non-native Americans dropping sounds in word initial consonant clusters.

Comparing assimilation rules in English and Chinese, English speakers sometimes use the extra linking sounds, *y*, *w* or *r*, for example: *She-y-isn't here*. *I'd like to-w-open the window*, and *America-r-and Canada* (Cleary et al. 2014) yet, they often go unnoticed.

When Bo used /y/ to link words in the phrase *What's-e 'yissue?* and Ying said, *Eh 'yif der are somting wrong,* it was also used between vowels, but was not a smooth link, occurred with a pause, and was too strongly stressed, which indicated the Chinese L1 preference to start words with a consonant.

Chinese is a syllable timed language, which means that every consonant-vowel syllable takes approximately the same amount of time, whereas English is a stress timed language with different stress given to the prominent syllables in the word and intonation group. Jenkins (2000) said that in isolation, stress is not a problem for comprehension, however, when the word *stability* was pronounced as /ˈstɪbɑːlɪdiː/ stressing the first instead of second syllable as well as changing two vowel pronunciations made comprehension very difficult.

To put this research into context, Table 8 shows a comparison with Jenkins (2000) Lingua Franca Core. Jenkins' LFC was written as a description of the areas speakers needed to be intelligible in a limited context. In this study, the focus was on what the listeners needed to be able to perceive what was said, also in a limited context. Both studies have a similar goal of mutual intelligibility, and it is important to note that the LFC is not used as a standard, but as a guide for critical pronunciation study.

Jenkins' LFC features necessary for speaker intelligibility	**Features necessary for recognition in this study**
No assimilation	No assimilation
Rhotic r	Rhotic r
Pronunciation of consonants	Pronunciation/perception of t, l, r, b, v & w

No deletion of consonants in word initial consonant clusters (in 2007, Jenkins also included word medial consonant clusters)	No deletion of consonants in any consonant clusters
Epenthesis is less likely to cause problems than elision, because the word is more recoverable.	No epenthesis (addition of extra sounds) between words
Vowel length, if stable, does not affect intelligibility	No vowel length variation
Vowel quality is not included because it is highly variable (in 2007 it was updated to being consistent or stable and not variable)	No vowel quality variation
Nuclear stress	Nuclear stress

Table 8: Comparison with Jenkins' Lingua Franca Core

O'Neal (2015) said that consonant pronunciation is consequential for maintaining intelligibility, including /z/ to /s/, /s/ to /θ/, and similarly in his 2014 paper on ELF interaction in Japan, he also found that consonant clusters in all positions are important for intelligibility. Jenkins (2007) clarified that recognition of weak vowels are important for receptive purposes. This research also showed that vowel quality was important for recognition. Similarly, vowel length was also important which is in disagreement with Jenkins and agreement with Deterding (2013). Having said that, Jenkins now puts a higher priority on negotiating and accommodating differences, which I agree is crucial in an active interaction. Assimilation by the speaker and for the listener makes intelligibility more difficult. Therefore, when students interact with many speakers who naturally use assimilation, they will need to learn how to decode or clarify what has been said. Therefore,

speakers, ELF or otherwise, who are considerate to non-native English listeners will dissimilate their words, pronounce their consonants, keep their vowel quality and quantity consistent, choose their vocabulary carefully, and pause at suitable junctures to allow the listeners some processing time. However, the listeners also need to learn about the problematic pronunciation features in table 7 to help them to recognize words, and develop accommodation strategies so they are able to be flexible when they don't understand by not being distracted by unknown phrases and also able to negotiate meaning by modifying problematic phonological segments to achieve mutual intelligibility.

Conclusion

Two important constraints, which made this an ESP context, were:

1) The presentations were at times, monologue, where comprehension was not acknowledged, and at other times, dialogue, during and after the presentation when there was a question, comment or discussion
2) The asymmetry and role of the participants limited their pragmatic choices, and ability to take turns

Although the listening processes may be similar between ESP, EFL and ELF contexts, the contextual factors are most certainly not. ELF contexts have different cultural, pragmatic and vocabulary norms from EFL contexts, as shown by the Asian Corpus of English (ACE) (Kirkpatrick 2010), and the Vienna-Oxford International Corpus of English (VOICE) (Seidlhofer 2001). However these corpora do not take into account the phonological

variations which can be the cause of communication breakdowns. Furthermore, in these video conferences, off-screen speakers and background noise, and restricted gestures, eye contact and non-verbal cues, all created additional ambiguity. This forced the listeners to rely more heavily on their bottom-up processing of the data stream.

The size of this data set means that it is just a pilot study and is not generalizable to a wider understanding of Asian BELF listening comprehension skills. However, it did enable the students to build awareness of the problematic pronunciation features they regularly encounter. It provided a clear pedagogical platform to build receptive and productive pronunciation skills, which is the first stage in developing their flexibility to accommodate variation. The main limitation of this pronunciation research was that the students did not show signs of acquisition of their new awareness; for example, the /v/ being pronounced as /w/ pattern was studied in the word *vital* in an earlier cycle. However, when it was heard again in the word *vinyl* in a later cycle, the students could not recognize it, even though it was a familiar word. Therefore, they were not able to use this knowledge to add to their decoding skills. In addition, there was no explicit assessment of listening skill development. The students were able to compare their transcripts with each other to understand their relative success rate, but there was not a categorical listening skill level evaluation, such as measuring the percentage of intelligible words (Munro & Derwing 1995). This limitation was because the students' chose the audio extracts based on a strong desire to understand the technical feedback about their contact lens research. In addition, they wanted to know why it was difficult to understand. So these two issues were at the expense of using standardized listening samples to control difficulty. This motivational

finding supports the need for authentic listening material development at the blurred boundary between listeners and users (Gao 2012).

Acknowledgements

I would like to thank George O`Neal, Anna Husson Isozaki, and the staff at the Japanese Association for Language Teaching Peer Support Group for their proof reading support.

References

Anderson, A. and Lynch T. (1988). *Listening.* Oxford: Oxford University Press.

Baker, W. and Jenkins J. (2015). Criticising ELF. *Journal of English as a Lingua Franca*; 4/1: 191-198

Baker, W. (2015). *Culture and identity through English as a lingua franca.* Berlin, Boston: De Gruyter Mouton.

Bargiela-Chiappini, F., Zhang Z. (2013) Business English. In B. Paltridge & S. Starfield (Eds.) *The Handbook of English for Specific Purposes* (pp. 193-212). Oxford, England: Wiley-Blackwell.

Basturkmen, H. (2010). *Developing Course in English for Specific Purposes.* Basingstoke, UK: Palgrave Macmillan.

Brown, G. (1990). *Listening to Spoken English*, London: Longman.

Cleary, C., Holden B. and Cooney, T. (2014). *Top-Up Listening 2* (2nd ed.) Tokyo: Abax

Cogo, A. (2009). Accommodating difference in ELF conversations. In A. Mauranen & E. Ranta (Eds.) *English as a Lingua Franca: Studies and Findings* (pp. 254-273). Newcastle, UK: Cambridge Scholars Press.

Cogo. A. (2015). Complexity, negotiability, and ideologies: a response to Zhu, Pitzl, and Kankaanranta et al. *Journal of English as a Lingua Franca* 4/1: 149-155.

Collins, B. and Mees, I. M. (2013). *Practical Phonetics and Phonology* (3rd Edition) Oxon, UK: Routledge

Conrad, L. (1983). Semantic versus syntactic cues in listening comprehension. *Studies in Second Language Acquisition* 7, 59-69.

Darwin, C. J., Turvey, M. T. and Crowder, R. G. (1972). An auditory analogue of the Sperling partial report procedure: Evidence for brief auditory storage. *Cognitive Psychology*, 3, 255-67.

Dauer, R. (2005) 'The Lingua Franca Core: A New Model for Pronunciation Instruction?' *TESOL Quarterly.* 39/3:543-660.

Davis, M. H. and Johnsrude, I. S. (2007). Hearing speech sounds: Top-down influences on the interface between audition and speech perception. *Hearing Research* 229: 132-47.
Deterding, D. (2013). *Misunderstandings in English as a Lingua Franca.* An Analysis of ELF Interactions in South-East Asia. Berlin: De Gruyter
Ehrenreich, S. (2009). English as a lingua franca in multinational corporations: Exploring business communities of practice. In A. Mauranen & E. Ranta (Eds.), *English as a lingua franca: Studies and findings* (pp. 126-151) Cambridge Scholars Publishing.
Ehrenreich., S. (2016). English as a lingua franca (ELF) in international business contexts: key issues and future perspectives. In K. Murata (Ed.) *Exploring ELF in Japanese Academic and Business Contexts.* (pp. 135-155). London and New York: Routledge.
Flowerdew J. (1996). *Academic Listening. Research Perspectives.* Cambridge: Cambridge University Press.
Field, J. (2004). An insight into listeners' problems: Too much bottom-up or too much top-down? *System* 32: 363-77.
Firth A. (1996). The discursive accomplishment of normality: On 'lingua franca' English and conversation analysis. *Journal of Pragmatics* 26: 237-259.
Gao, Y. (2012). University students' motivation on learning English and their identity development: A four year longitudinal study, Beijing: Higher Education Press.
Gilbert, J. & Levis, J. (2001). Review of The Phonology of English as an International Language. *TESOL Quarterly* 35/3. 505-506.
Goh, C. C. M. (1998). How ESL learners with different listening abilities use comprehension strategies and tactics. *Language Teaching Research*, 2/2, 124-147.
Goh, C. C. M. (2000). A cognitive perspective on language listeners' listening comprehension problems. *System* 28: 55-75.
Grazzi, E. (2003). Review of The Phonology of English as an International Language. Perspectives. *A Journal of TESOL* – Italy. 30/2.
Guion, S.G., Flege, J.E., Akahane-Yamada, R., Pruitt, J.C., (2000). An investigation of current models of second language speech perception: the case of Japanese adults' perception of English consonants. *J. Acoust. Soc. Am.* 107, 2711–2724.
Gudykunst W.B. & Ting-Toomey S. (1988). *Culture and Interpersonal Communication.* Newbury Park: Sage.
Hall E.T. (1976). *Beyond Culture.* New York: Doubleday.
Hallé, P. A.; Best, C. T.; Levitt, A. (1999). Phonetic vs. phonological influences on French listeners' perception of American English approximants, *Journal of Phonetics*, 27: 281–306.
Hewings, M. (2001). Review of The Phonology of English as an International Language. ELT Journal 55/3, 327-329.
Hofstede G. (1983a). Dimensions of national cultures in fifty countries and three regions. In J. Deregowski, S. Dzuirawiec & R. Annis (Eds) *Explications in Cross-cultural psychology.* Lisse: Swets & Zeitlinger.
House J. (2003). English as a lingua franca: A threat to multilingualism*? Journal of Sociolinguistics 7*: 556-578.
Jenkins, J. (2000). *The Phonology of English as an International Language: New Models, New Norms, New Goals.* Oxford: Oxford University Press.

Jenkins, J. (2007). *English as a lingua franca: Attitude and Identity.* Oxford; New York: Oxford University Press.
Jenkins, J. (2008) Misinterpretation, Bias and Resistance to Change: The Case of the Lingua Franca Core. In Dziubalska-Kołaczyk K. and Przedlacka J. (Eds.), *English Pronunciation Models: A Changing Scene.* (2nd edition.). (pp. 199-212). Bern – Frankfurt am Main: Peter Lang.
Jenkins, J. (2015). Repositioning English and multiculturalism in English as a Lingua Franca. *English in Practice* 2(3): 49-85.
Kachru, B. (1985). Standards, codification and sociolinguistic realism: The English language in the outer circle. In R. Quirk & H. Widdowson (Eds.) *English in the World: Teaching and Learning the Language and Literatures* (pp.11-30). Cambridge: Cambridge University Press.
Kachru, B. B. (1992). Models for non-native Englishes In B. B. Kachru (Ed.) *The Other Tongue: English Across Cultures* (pp. 48-74). Urbana, IL: University of Illinois Press.
Kachru, B. B. (2005). *Asian Englishes: Beyond the Canon.* Hong Kong: Hong Kong University Press.
Kirkpatrick (2010). *English as a Lingua Franca in ASEAN: A multilingual Model.* Hong Kong: Hong Kong University Press.
Louhiala-Salminen, L., Charles, M. & Kankaanranta, A. (2005) English as a lingua franca in Nordic corporate mergers: Two case companies. *English for Specific Purposes* (Special issue: English as a lingua franca in international business contexts), C. Nickerson (ed.), 24(4), 401-421. DOI: 10.1016/j.esp.2005.02.003
Lynch T. (1996). *Communication in the Language Classroom.* Oxford: Oxford University Press.
Mann, S. (2011). A Critical Review of Qualitative Interviews in Applied Linguistics. *Applied Linguistics* 32 (1): 6-24.
Matsumoto, Y. (2011). Successful ELF communications and implications for ELT: Sequential analysis of ELF pronunciation negotiation strategies. *The Modern Language Journal* 95: 97-104.
Mauranen, A. (2006). Signaling and preventing misunderstanding in English as lingua franca communication. *International Journal of the Sociology of Language*, 177, 123-150.
Munro, M. and Derwing, T. (1995). Foreign Accent, Comprehensibility, and Intelligibility in the Speech of Second Language Learners. *Language Learning* 45/1.
Munro, M. and Derwing, T., & Morton, S. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition*, 28, 111-131.
Nelsen, C. L. (2011). *Intelligibility in World Englishes – Theory and Application.* London/New York: Routledge.
O'Neal, G. (2014). The Link between Pragmatics and Phonology. *Pragmatic Matters* 43
O'Neal, G. (2015). Interactional intelligibility: the relationship between consonant modification and pronunciation intelligibility in English as a Lingua Franca in Japan. Asian Englishes, 17 (3), 222-239.
O'Regan, J. (2014). English as a lingua franca: An immanent critique. *Applied Linguistics* 35(5), 533-552.

Pow, B. (2002). Review of The Phonology of English as an International Language. BRAZ-TESOL Newsletter, March 2002/12.
Qian, W. (2011). Phonological Features of China English: An Acoustic Investigation on Segmental Features of Educated China English Speakers. *Proceedings of the 16th Conference of Pan Pacific Association of Applied Linguistics.*
Pickering, L. (2006). Current Research on Intelligibility in English as a Lingua Franca. *Annual Review of Applied Linguistics* 26, 219-233.
Pinker, S. (1994). *The Language Instinct.* London: Penguin.
Rajadurai, J. (2007). Ideology and Intelligibility, Intelligibility studies: a consideration of empirical and ideological issues. *World Englishes* 26 (1): 87-98.
Renandya, W. A. and Farrell, T. S. C. (2011) 'Teacher, the tape is too fast!' Extensive listening in ELT, *ELT Journal*. 65/1: 52-59.
Richards, J. C. (1983). Listening Comprehension: Approach, design, procedure. *TESOL Quarterly*. 17: 219-40.
Rivers, W. M. (1981). *Teaching foreign language skills.* Second Edition. Chicago: University of Chicago Press.
Rogerson-Revell, P. (2007). Using English in international business: A European case study. *English for Specific Purposes* 26: 103-20.
Schwartz, G. (2008). The Lingua Franca Core and the Phonetics-Phonology Interface. In Dziubalska-Kołaczyk K. and Przedlacka J. (Eds.), *English Pronunciation Models: A Changing Scene.* (2nd edition.). (pp. 131-150). Bern – Frankfurt am Main: Peter Lang.
Shaw, P., Caudery, T. and Petersen, M. (2009). Students on exchange in Scandinavia: Motivation, interaction, ELF development. In A. Mauranen & E. Ranta (Eds.) *English as a Lingua Franca: Studies and Findings* (pp. 178-199). Newcastle, UK: Cambridge Scholars Press.
Simpson, A. (2014). Business English as a lingua franca: Repair, preference, and turn taking. In N. Sonda & A. Krause (Eds.), *JALT2013 Conference Proceedings.* Tokyo: JALT.
Seidlhofer, B. (2001). Closing the conceptual gap: The case for a description of English as a Lingua Franca. *International Journal of Applied Linguistics.* 11 (2): 133-158.
Seidlhofer, B. (2008). ELF findings: form and function. Plenary paper given at the First International ELF Conference, Helsinki, March 2008.
Seidlhofer, B. (2009). Orientations in ELF research: form and function. In A. Mauranen & E. Ranta (eds.) 2009.
Serafini. E., J., Lake., J., B., Long. & M., H. (2015). Needs analysis for specialized learner populations: Essential methodological improvements. *English for Specific Purposes.* 40: 11-26.
Smith, L., E., and Bisazza, J., A. (1982). The comprehension of three varieties of English for college students in seven countries. *Language Learning* 32: 259-270.
Smith, L., E., and Rafiqzad, K. (1979). English for cross-cultural communication: The question of intelligibility. *TESOL Quarterly* 13: 371-180.
Sobkowiak, W. (2008). Why Not LFC? In Dziubalska-Kołaczyk K. and Przedlacka J. (Eds.), *English Pronunciation Models: A Changing Scene.* (2nd edition.). (pp. 131-150). Bern – Frankfurt am Main: Peter Lang.

Sowden, C. (2012). ELF on a mushroom: the overnight growth in English as a Lingua Franca. *ELT Journal* 66 (1), 89-96.
Schwartz, G. (2008). The Lingua Franca Core and the Phonetics-Phonology Interface. In Dziubalska-Kołaczyk K. and Przedlacka J. (Eds.), *English Pronunciation Models: A Changing Scene.* (2nd edition.). (pp. 177-198). Bern – Frankfurt am Main: Peter Lang.
ten Have, P. (1999). *Doing Conversation Analysis: A Practical Guide*. London: Sage.
Vandergrift, L. (1998a). Constructing meaning in L2 listening: Evidence from protocols. In Lapkin S. (Ed.), *French second language education in Canada: Empirical studies* (pp. 89–119). Toronto: University of Toronto Press.
Walker, R. (2001). Pronunciation for international intelligibility. *English Teaching Professional, 21*: 1-7.
Wang, Y. and Jenkins, J. (2016) "Nativeness" and intelligibility: Impacts on intercultural experience through English as a lingua franca on Chinese speakers' language attitudes. *Chinese Journal of Applied Linguistics*, 39, (1), 38-58. (doi:10.1515/cjal-2016-0003).
Zhang, F. and Yin P. (2009). A Study of Pronunciation Problems of English Learners in China, *Asian Social Science,* 5/6 Retrieved on January 2 2015, from http://www.ccsenet.org/journal/index.php/ass/article/viewFile/2490/2336
Zhu H. (2015). Negotiation as the way of engagement in intercultural and lingua franca communication: frames of reference and interculturality. *Journal of English as a Lingua Franca* 4/1: 63-90.
Zoghbor. W. S. (2011). Teaching the Pronunciation of English as a Lingua Franca: Reducing Skepticism and Increasing Practicality. *International Journal of Humanities and Social Science.* 1/8: 285-288.
Zuber-Skerritt., O. D. (2001). Action learning and action research: paradigm, praxis and programs. In S. Sankaran, B. Dick. R. Passfield, Swepson P., (Eds.) (pp. 1-20). Lismore NSW, Aust.: Southern Cross University Press

Appendix – Transcription Conventions

(2.0)	Pause of about 2 second	(..)	Pause of about 0.5 second
[	Overlap	[[	Speakers start at the same time
= =	Latched utterances	____	Emphasis
-	Abrupt Cut-off	::	Sound stretching
(xxx)	Unable to transcribe	° °	Quiet
( )	Unclear word or phrase	(( ))	Comment or non-speech activity
>word<	Quicker speech	<word>	Slower speech
↑	Rising Intonation	↓	Falling Intonation
//	Phonemic transcription	[ ]	Phonetic transcription