University of Reading

EFL Teachers' Attitudes to Intelligibility and Nativeness in Pronunciation Assessment

Doctor of Philosophy

Department of English Language and Applied Linguistics, University of Reading

Ibrahim Mohammed Hussain Alfaifi

March 2022

Abstract

This study has explored non-native English-speaking (NNS) and native English-speaking (NS) EFL teachers' attitudes to intelligibility and nativeness in pronunciation assessment. The study is timely and relevant due to the importance of investigating EFL teachers' language attitudes as highlighted in previous literature (e.g., Dyers & Abongdia, 2010). It is also relevant since its findings can help in identifying effective ways for developing professional training for teachers on assessing pronunciation and, by extension, teaching pronunciation, and contribute to recent research on best practice regarding pronunciation assessment and, by extension, pronunciation instruction. Furthermore, this study is needed due to the limited research in EFL teachers' attitudes research demonstrated in previous literature (Baker, 2014; Baker & Murphy, 2011), especially in pronunciation assessment. The study has also particular relevance to its context, King Abdulaziz University (KAU), where the author of this study is an EFL teacher. At the university there are teachers of English from all over the world and many of these teachers do not support the use of norms based on nativeness, as this study demonstrates.

The current study adopted a mixed-methods design; it made use of both quantitative and qualitative methods, and direct and indirect methods to answer the research questions. It elicited the attitudes of the teachers through two practical tasks and one online questionnaire. The practical tasks included an assessment task that involved rating L2 English speakers on phonological features and an acceptability task that involved responding to spoken words by accepting or rejecting them. The online questionnaire was an agreement/disagreement questionnaire based on a 7-point scale. The study tasks were followed by interviews with a small number of the study participants (N = 6) to further investigate the behaviour of these participants in the study tasks. The participants of the study are EFL teachers (N = 57). A questionnaire was used to elicit personal background information about the participants such as their ages, teaching experience and speaking assessment experience.

The study adopted a quantitative approach to data analysis. Percentages were calculated for the data of the assessment of students' speech task and then comparisons between the participants were conducted using t-tests to determine the attitudes of the

participants towards intelligibility and nativeness. Also, comparisons for the data of the acceptability task and the online questionnaire were made using t-tests to determine the attitudes of the participants towards intelligibility and nativeness. Next, correlation analyses were conducted for the data of the practical tasks (i.e., the assessment of students' speech task and the acceptability task) and the data of the online questionnaire to find out if the attitudes in the online questionnaire are reflected in the practical tasks. Further, correlation analyses were made to investigate if the attitudes of the participants in the study tasks are affected by particular personal background information such as age and teaching experience. Finally, deductive analysis for the data of the follow-up interviews was conducted, where the research questions were determined as categories and then any connections in the data of the follow-up interviews were mapped to these categories.

The findings of this study showed that both groups of participants (NNS and NS) have behaved similarly in the practical tasks: neither were intelligibility or nativeness. However, in the online questionnaire, whereas the NNS group failed to show more orientation to intelligibility or nativeness, the NS group were more oriented to intelligibility. The findings also showed that the attitudes of the NS group were more oriented to intelligibility than the attitudes of the NNS group in the online questionnaire, but both groups were similar as regards nativeness. The findings further showed that the participants' attitudes in the online questionnaire may have not been reflected in their behaviour in the practical tasks. The attitudes of the participants emerged in the online questionnaire were not correlated with their behaviour in the practical tasks. The reasons behind this according to the interviews conducted may have been that the participants' real attitudes in the online questionnaire were hidden and their behaviour in the practical tasks were the representative of their attitudes. Finally, the findings showed that in general the participants' age, teaching experience and speaking assessment experience had no effect on the participants' attitudes. However, according to the interviews conducted, experience as L2 English learners at school, experience of learning English in general and speaking assessment policies were possible factors affecting the participants' attitudes.

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Candidate:
Ibrahim Alfaifi

Table of Contents

List of Tables

List of Figures

CHAPTER 1: INTRODUCTION

1.1 Introduction

This study has explored non-native speaker (NNS) and native speaker (NS) EFL teachers' attitudes to intelligibility and nativeness in pronunciation assessment. It elicits the attitudes of the teachers through three tasks: (1) an agreement/disagreement questionnaire based on a 7-point scale; (2) an assessment task that involves rating L2 English speakers' pronunciation; and (3) an acceptability task that involves responding to spoken words and phrases by accepting or rejecting them. The study adopts the Lingua Franca Core (LFC; Jenkins, 2000) as a framework for intelligibility. The LFC is a description for the phonological features essential for intelligibility in English as a lingua franca (ELF; see section 3.5, Chapter 3).

This study uses the terms English NS and NNS to distinguish between the English L1 and L2 users. Some scholars perceive the use of NNS vs NS as a problem and argue for the use of an L1 and L2 user, while some others see the use of an L1 vs L2 user as problematic and argue for the other use (see e.g., Dewaele, 2018). This study is not meant to address this line of discussion. The study's main scope is to investigate the attitudes of both kinds of speakers and to find out if they adopt different attitudes. However, the definitions of both NS and NNS and the differences between both speakers are introduced below.

Classifying the speakers of English as NNS and NS is a thorny issue (Davies, 2003). The models that classified speakers of English as NNS and NS did not provide specific definitions of those two kinds of speakers (ibid.). Davies (ibid.) has widely studied the issue of NS and thus he suggested that it is difficult to define a NS given the diverse aspects that need to be considered when defining a NS, such as psycholinguistic, linguistic and sociolinguistic perspectives (ibid.). Davies (ibid.) pointed out that the concept "native speaker" has been used by applied linguistics and foreign language teachers as a common reference point, although this concept is highly ambiguous, and it is difficult to define who exactly a native speaker is. The use of this concept could also cause racism as it could exclude speakers of certain native varieties or very proficient NNSs (Dewaele, 2018). This concept is also misleading as it suggests a very high level of proficiency of first language (ibid.). First language proficiency can be lost gradually by acquiring a second language (Schmid, 2013), which necessarily changes the status of a person as a native speaker (Dewaele, 2018).

Davies (1991) defined a NS as someone has a high degree of competence in a language and linguistic intuition of that particular language. Tay (as cited in Davies, 2003) provided another definition of a NS: "One who learns English in childhood and continues to use it as his dominant language, and has reached a certain level of fluency". Tay (ibid.) further argued that there are three conditions when referring to someone as a NS: "If a person learns English late in life, he is unlikely to attain native fluency in it; if he learns it as a child, but does not use it as his dominant language in adult life, his native fluency in the language is also questionable; if he is fluent in the language, he is more likely one who has learned it as a child (not necessarily before the age of formal education but soon after, and has continued to use it as his dominant language".

Davies (2003) argued that there are two perspectives regarding a NS: as a myth and as a reality. He explained that the reality of a NS has to do with his/her "membership" (ibid., p. 7). He stated that a NS is relied on to know how linguistic things are done since s/he is the repository of the language (ibid.). A NS is also expected to have fluent connected speech, and to have control over expected characteristic strategies of performance and communication (ibid.). Given the NS's intuitive feeling and their "characteristic systematic set of indicators" such as linguistics, pragmatics, paralinguistics, and shared cultural knowledge (ibid., p. 207), a NS can identify other members of NSs' community (ibid.). A NS as a myth is perceived as an idealized speaker (ibid.). Davies (ibid.) argued that a NS as a myth has the following characteristics:

- A NS acquires his/her first/native language in childhood;
- A NS has intuitions regarding acceptability and productiveness of his/her grammar;
- A NS has intuitions regarding the standard grammar features which are different to those of his/her idiolectal grammar;
- A NS is able to produce fluent spontaneous discourse, facilitated by memorized lexical items;
- A NS has a creative ability which enables him/her to produce creative discourse; and
- A NS is able to translate into his/her other first/native languages.

The opposite of the concept "native speaker" is the concept "non-native speaker. Dewaele, 2018) argued that the concept "non-native speaker" is inherently strange as we do not usually

define somebody by what s/he is not. He argued that we do not define, for example, blue-eyed people as not brown-eyed (ibid.).

A NNS was defined as someone who knows and uses a second language at any level (Cook, 2007). Cook (ibid.) explained that since second language acquisition (SLA) research and language teaching have considered a NS as the standard starting point, a NNS is considered to deviate from NSs' norms. However, Cook (ibid.) argued that a NNS needs to be considered a speaker different from a NS rather than as a deficient speaker. He further argued that a NNS has particular characteristics (ibid.): (a) A NNS has uses of his/her second language different to those of a NS; (b) a NNS usually needs to adapt and change his/her language competence in to order to effectively communicate with a NS; (c) a NNS usually needs to translate information to understand a NS; and (d) a NNS attends to two languages (i.e., first and second languages) at once while using a second language. Cook (ibid.) stated that the knowledge of a NNS of a second language is different to that of a NS. Spelling, writing and grammar are some of the language skills in which a NNS may have a more advanced level than a NS, while a NS may be more advanced in oral skills than a NNS (ibid.). Cook (ibid.) further explained that a NNS may have a different mind to a NS's. A NNS's interior aspects of mind go beyond external uses of language. One of these aspects is NNS's awareness when referring to words, a sentence structure and oral and cultural aspects of the second language (ibid.).

1.2 Rationale of Study

This study is relevant due to the importance of investigating EFL teachers' language attitudes. According to Dyers and Abongdia (2010), language attitudes are "what people think about a language, how a language makes them feel, and what they are prepared to do about that language" (p. 121). Given this concept of language attitudes, investigating EFL teachers' language attitudes is important as EFL teachers' language attitudes can affect the way EFL teachers shape learning experience, and sometimes their attitudes can affect the way they set teaching outcomes and assessment criteria if they are responsible for doing so, like the case at King Abdulaziz University (KAU, as far as the researcher is aware). For example, if the attitudes of an EFL teacher were oriented to nativeness, s/he would focus on teaching his/her students to sound nativelike and s/he would use teaching and assessment materials and criteria based on nativeness. Such a teacher would motivate his/her students to learn how to sound nativelike, and s/he would set nativeness as a criterion for successful teaching and learning outcome. S/he would also set nativeness as a criterion when s/he assess how a student is successful in learning L2 English, if s/he was

responsible for assessment criteria. In return, learning experience and assessment can affect learners' language attitudes (Monfared & Khatib, 2018). Monfared and Khatib (ibid.) further explained that teachers' practices can have a long-term impact on learners' language attitudes and preference owing to the learners' lack of professional knowledge. In return, those learners' language attitudes may persist to interfere in their future teaching career creating a series of similar attitudes from a generation to another. This was notified by Graves (2000) as she stressed the importance of learning experience. Graves (ibid.) referred to the effect of teachers' experience as learners on their beliefs, and thus inevitably their practices in the classroom (ibid.). This was echoed by Brown and McGannon (2007, as cited by Monfared & Khatib, 2018); they highlighted those false perceptions that teachers have about L2 learning (e.g., blind followers of nativeness) can affect their teaching practices (ibid.), and therefore their learners' perceptions (ibid.).

Also, this study is relevant since it helps in developing teachers' training and inform future research. According to Gurzynski–Weiss (2013), researchers within the area of instructed second language acquisition (SLA) are directing their attention to a great variety of individual differences among language instructors, including the theoretical and practical knowledge the teachers bring along to the teaching context, which includes their perceptions (ibid.). Those researchers (e.g., Basturkmen, 2012) are trying to identify whether teachers' theoretical and practical knowledge are complying with the researchers' conclusions and pedagogical specialists' recommendations of best practices so they can identify effective ways of developing appropriate professional training for teachers (Nagle et al., 2018). Also, this kind of investigation could inform further kinds of research in the future. For example, according to Robinson (1996), researchers often rely on experienced teachers' beliefs and knowledge for tackling some complex issues in L2 pedagogy and assessment, such as examining grammatical features that are tricky to certain EFL learners (ibid.). Given all this, the current study becomes particularly relevant as investigating EFL teachers' attitudes in pronunciation assessment can (a) help identify effective ways for developing professional training for teachers on assessing pronunciation, and by extension teaching pronunciation; and (b) inform future research, e.g., tackling complex issues in pronunciation assessment and instruction.

Furthermore, this study is needed due to the limited research in EFL teachers' attitudes research, especially in pronunciation (Baker, 2014; Baker & Murphy, 2011). According to Buss (2016), only in recent years we have seen some research done on L2 teachers' cognition regarding L2 pronunciation in different ESL and EFL contexts. Also, Kubanyiova and Feryok (2015) suggested that we still do not have enough knowledge as regards language teacher cognition on L2

pronunciation and how it can be developed. Furthermore, while only limited research has been done on L2 teachers' cognition regarding L2 pronunciation, especially in attitudes to nativeness and intelligibility in pronunciation among L2 English teachers (see Chapter 4 for such research), there is particularly a lack of research of this type in the field of pronunciation assessment. In this area, among L2 English teachers, there is only one study (Bøhn & Hansen, 2017), as far as the author is aware. However, Bøhn and Hansen (ibid.) collected only teacher-examiners' opinions regarding intelligibility and nativeness in pronunciation assessment for their study data, i.e., they did not present any evidence of the teachers' attitudes in practical tasks (see section 4.2.2.2, Chapter 4, for more details on this study). Thus, more research addressing such aspects is required to find out if EFL teachers' attitudes are translated into their practices, as attitudes are often hidden behind a social facade (Dalton-Puffer, Kaltenbock, & Smit, 1995; Giles & Billing, 2004).

The study has also particular relevance to its context, KAU, where the author of this study is an EFL teacher. The students take in their foundation year at KAU an EFL course before they begin their studies, since English is the medium of instruction at the university for most programmes. Based on their scores on this course, the students are admitted into different programmes. Thus, the EFL assessments at the university, which include pronunciation assessment, are of major importance. Consequently, it is important to know the attitudes of the EFL teachers at KAU as they are responsible for making and conducting the EFL assessments. The teachers' attitudes would reflect on the assessments and then on the students' attitudes. They would also reflect on the attitudes of the students when they become EFL teachers in the future (as mentioned above). Thus, it is important to find out if the EFL teachers' attitudes are oriented to intelligibility in pronunciation assessment more than nativeness, given the importance of intelligibility over nativeness demonstrated in the literature (see Chapters 2 and 3). If they are not, there will be a need for training for these teachers in this regard. Overall, the study would also be relevant to many other contexts with similar conditions as those at KAU.

1.3 Context of Study

Most of the participants in this study are EFL instructors for the Preparatory Year English Language Program at King Abdulaziz University (KAU), Saudi Arabia. The following subsections introduce information about this programme. All this information is retrieved from KAU's website for its English Language Institute (2017).

1.3.1 Preparatory Year English Language Program at KAU

The Preparatory Year English Language Program at KAU is a compulsory component of the university Preparatory Year Program for all newly admitted full-time KAU students. The program consists of two tracks: General English Track (GET), which is for students whose majors use Arabic as the medium for instruction; and Academic English Track (AET), which is for students whose majors use English as the medium for instruction. Both tracks are designed to equip students with intermediate language proficiency level in using the English language, equivalent to the B1 band set by the Common European Framework of Reference (CEFR). For the GET students, this is equivalent to almost the score of 5 or above in General International English Language Testing System (IELTS) or the score of 45 or above in internet-based Test of English as a Foreign Language (iBT TOFEL). For the AET students, this is equivalent to almost the score of 5.5. in the academic IELTS or the score of 59 in the iBT TOFEL, which is the minimum level of the English language proficiency level accepted for the students to be admitted into the KAU schools, following the Preparatory Year Program.

1.3.2 Program Course Materials

All course materials in the program are mapped onto the proficiency level descriptors of the CEFR. The program is run using a modular system, where there are four modules, two modules each academic term. The duration of each module is seven academic weeks, with 18 hours long of instruction per week. In partnership with Cambridge University Press (CUP), the program uses Unlimited Special Edition (EUSE) instructional materials for the GET and the Unlock series of materials for the AET. The GET materials are aligned with the CEFR and thus they are designed to help students proceed from the Beginner Level (CEFR A1 level) to the Intermediate Level (CEFR B1+) in one academic year. The AET materials are also mapped onto the CEFR and designed to assist students in progressing from CEFR A1 level to CEFR B1 level with more attention given to academic skills and academic vocabulary.

1.3.3 Program Speaking Assessment

Based on the students' scores on the program assessment, the students are admitted into different majors at the university. This assessment involves the assessment of speaking which includes the assessment of pronunciation together with fluency. The speaking assessment for KAU's Preparatory Year English Language Program is conducted by two teachers. One teacher acts as an interlocutor and grader, and the other teacher acts as a grader and observer only. The speaking

assessment is done through two speaking tasks. The first task is short interview questions, where the students are asked five short questions that require short answers from the students. The second task is a short monologue task with further questions and answers. In this task the students take two minutes prior to the task to prepare themselves and then they speak about a topic and points they are given during the preparation period. This is followed by questions from another student who represents the audience, and the speaker should give short answers to these questions (see Appendix A for the original assessment guidelines used for the program).

1.3.4 Program Speaking Assessment Scale

The program speaking assessment is conducted using a scale constructed by some of the EFL teachers at the university. They have based the descriptors of this scale on the following sources: speaking proficiency scales of CEFR A2 level at St. Giles International (CEFR Level A2, n.d.), and the oral assessment criteria of CEFR A2+ level in the Council of Europe Manual: "Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR)" (Table C3; Figueras et al., p. 186, 2009). In the following two paragraphs, details about A2 level in the first source and A2+ level in the latter source are presented. These details were drawn from the two sources.

The students at the A2 level can easily use phrases on familiar topics in short communications, but with obvious hesitation and false starts. The students can also use frequent expressions to handle basic language functions: e.g., speaking about personal details, daily routines, desires and requirements, or requests for information. Furthermore, the students at the A2 level can construct simple sentence patterns and use memorised expressions in short exchanges. They can also use groups of a few words to talk about basic things: e.g., themselves and other people, what they do, places or possessions. However, the memorised phrases the students use are short, limited and covering predictable situations. In non-routine situations, they have recurrent breakdowns during communication along with frequent misunderstandings. The students at the A2 level have a good repertoire of vocabulary for handling basic communicative requirements and for surviving in simple conversational exchanges. They can also handle everyday basic exchanges and use simple sentence structures accurately, but they still commit basic mistakes such as tense confusion and not marking agreement (for example, with a third-person singular subject (e.g., *John*), the verb does not have the *-s* suffix ending). However, the message they are trying to convey is usually delivered. The pronunciation of the students at the A2 level is normally clear enough to be understood despite an

obvious foreign accent, but the listener needs occasional repetitions before s/he can understand the speaker of this level. The students at the A2 level can also combine groups of words with basic connectors such as “and”, “but” and “because”. They can further make brief social conversations using common forms of greeting and address. They can make and respond to apologies, invitations, and suggestions, and they can ask for attention and construct a straightforward narrative.

The students at A2+ level have a good repertoire of vocabulary that enables them to carry out everyday communications in familiar situations. However, due to their limited knowledge of vocabulary, they usually find it difficult to deliver a complete message and they hesitate while looking for words that express their meanings. The students at A2+ level can also use memorised simple phrases to easily survive in brief everyday communications, despite obvious hesitation and false starts. They can also start, keep and close brief simple conversations. They can also ask and answer questions on topics of interest, hobbies and past activities. However, they find it difficult to participate in open discussions. The students at A2+ level can also use the most recurrent connectors to link simple sentences in order to tell a short story or construct a straightforward narrative.

In the scale used in KAU’s Preparatory Year English Language Program, there are five levels of descriptors that describe different proficiency levels/scores, and thus for both pronunciation and fluency there are five levels of descriptors. A summary of the five levels/scores and their descriptors are given in the table 1.1. below (see Appendix B for the original speaking rating scale used in KAU’s Preparatory Year English Language Program).

As noted from the table 1.1 below, the descriptors are holistic and vague when it comes to describing the proficiency levels of pronunciation, and sometimes irrelevant to both pronunciation and fluency. For instance, for level 5 it says, “student’s pronunciation is reasonably accurate for main vocabulary related to predictable situations” and “can handle short, routine exchanges with relative ease”. For level 3 it says, “short utterances, pauses, false starts, and reformulation are common”. For level 0 it says, “the pronunciation is completely unintelligible”. Some of the descriptors are also oriented to native-speaker norms; they use nativeness, or sounding nativelike, as a criterion for L2 English pronunciation proficiency. For example, in level 3 it says, “student can make him/herself understood despite a foreign accent”. Also, for level 2 it says, “pronunciation of a very limited repertoire of learnt words and phrases can be understood with some effort by native speakers used to dealing with Arabic speakers”. In these examples a foreign accent, including an

Arabic English accent, was considered problematic for pronunciation proficiency. Also, native speakers were considered as an authority over what is intelligible as their pronunciation was assumed to be the acceptable norms. Thus, in these examples the scale seems to consider nativeness as a criterion for pronunciation proficiency.

Table 1.1: Speaking rating scale of KAU's Preparatory Year English Language Program

Score	Descriptors (including pronunciation and fluency)
5	• Student has a good repertoire of vocabulary so that s/he can construct basic sentence patterns reasonably accurately but with some searching for words. • Student can interact with reasonable ease. S/he can ask easily understandable questions and answer questions adequately. • Student's pronunciation is generally correct for main vocabulary. • Hesitations and false starts may still be quite obvious.
4	• Student has a good repertoire of vocabulary so that s/he can construct basic sentence patterns but with some searching for words and some basic grammatical mistakes. • Student can ask and answer questions using basic sentence structures. • Student's pronunciation is generally correct for main vocabulary. • Hesitations and false starts may still be quite obvious.
3	• Student can answer questions using basic sentence structures, memorised phrases, and groups of few words, with some basic grammatical mistakes. • Student is understood despite a foreign accent. • Short utterances, pauses, false starts and reformulation of phrases are frequent. • Student may not be understood from the first time.
2	• Student has a limited repertoire of vocabulary so that s/he can only construct basic phrases but with some basic grammatical mistakes. • The use of short, memorised phrases is frequent with much pausing. • Communication is done through repetition, reformulation of phrases and false starts. • Student is understood with some effort by native speakers familiar with Arabic speakers.
1	• Student has very limited repertoire of vocabulary so that s/he can only construct basic phrases. • Student has difficulty interacting despite the use of repetition, reformulation of phrases, pausing and false starts. • Student is understood but with considerable effort by native speakers familiar with Arabic speakers.
0	• Student is absent, his/her answers are irrelevant to the prompts, or his/her pronunciation is totally unintelligible.

1.4 Outline of Thesis

This chapter has introduced the study by firstly providing an overview of the research. This was followed by a description of the study context. Finally, this chapter has discussed the rationale behind this study.

Chapter 2 introduces a literature review on pronunciation assessment. It firstly gives an overview on pronunciation assessment and then addresses the history of pronunciation assessment. It then presents two contradictory principles which have been long discussed in L2 pronunciation: intelligibility and nativeness. The chapter defines both principles and explains the shift of priority from nativeness to intelligibility. Next, the chapter discusses the linguistic and speech features essential for intelligibility. It then discusses pronunciation assessment scales. Finally, it concludes with a discussion on factors affecting pronunciation assessment.

Chapter 3 introduces another literature review; it introduces a lingua franca approach to intelligibility. It firstly gives an overview on English as a lingua franca. It then discusses the pedagogy, curriculum and culture of English as lingua franca. The chapter then discuses teacher education regarding English as lingua franca. Next, the chapter addresses language assessment and L2 English pronunciation in light of English as lingua franca. It then introduces a lingua franca approach to L2 English intelligibility, particularly L2 English pronunciation intelligibility. Finally, the chapter introduces a framework for L2 English pronunciation intelligibility based on a lingua franca approach to L2 English intelligibility: the Lingua Franca Core (LFC).

Chapter 4 introduces a literature review on language attitudes. It firstly gives an overview on language attitudes, and then it defines language attitudes research and discusses the popular methodologies used in language attitudes research. Finally, the chapter presents the literature on EFL teachers' language attitudes and towards the end it introduces the research questionnaire for this exploratory study.

Chapter 5 presents the methodology used in this study. After introducing the research questions again, the chapter explains the research design. It then describes the ethical consideration and the participants of this study. Next, the chapter explains the instruments used in this study to collect the data. This is followed by the procedures used to collect the data. Finally, the chapter describes the piloting process of the instruments and then the data analysis.

Chapter 6 presents the results of this study. It firstly presents the results of the study tasks. Next, the chapter presents the findings of the correlation analyses conducted for the dependent variables of this study and the correlation analyses conducted for the independent variables of this study and the dependent variables. Finally, the chapter presents the results of the follow-up interviews.

Chapter 7 is about the discussion of the study results. This chapter begins with a summary of the answers to the research questions. It then provides more details of the answers to the research questions through the chapter and discusses the study results in light of literature. Finally, the chapter discusses the limitations of the study.

Finally, Chapter 8 is the conclusion where this chapter closes this exploratory study. This chapter begins with a summary of the main findings of this study. It then discusses the contributions of this study to the literature. It finally discusses the implications of this study and makes recommendations for further research.

CHAPTER 2: LITERATURE REVIEW 1: PRONUNCIATION ASSESSMENT

2.1 Introduction

Since this PhD study looks at the attitudes of EFL teachers towards intelligibility and nativeness in pronunciation assessment, this chapter includes a background on pronunciation assessment. It first provides an overview on pronunciation assessment, and then it addresses the history of pronunciation assessment including the history of its neglect and resurgence. It then discusses two main contradictory criteria/models/principles which are often discussed in L2 English pronunciation teaching: nativeness and intelligibility. Given that these two principles are a main aspect of investigation in this PhD study, the chapter introduces definitions of these two principles, along with arguments and empirical research justifying them. It also explains the shift of focus from nativeness to intelligibility as models or criteria for L2 English pronunciation instruction. The explanation of this shift is important to understand the rationale behind the focus of this PhD study on the attitudes towards nativeness and intelligibility. This PhD study looks at whether EFL teachers' attitudes are consistent with recent research and literature in terms of a move away from nativeness and towards intelligibility. Following this, the chapter discusses the speech features essential for intelligibility. This section of the chapter is essential since it introduces the framework of intelligibility used in this PhD study: the LFC. Since this study has investigated the effect of pronunciation assessment policies on EFL teachers' attitudes, the current chapter discusses pronunciation assessment scales. Finally, the chapter concludes with a discussion on factors affecting pronunciation assessment.

This chapter refers to L2 English pronunciation pedagogy from time to time since both pronunciation assessment and pedagogy share many characteristics and the literature on pronunciation assessment per se is limited. Another reason is that some researchers have discussed L2 English pronunciation in terms of L2 English pronunciation pedagogy only, while pronunciation assessment is not always independent from L2 English pronunciation pedagogy. For example, the models "nativeness" and "intelligibility" cannot be discussed from the viewpoint of L2 English pronunciation pedagogy research without the ability to extend this discussion to pronunciation assessment research, and vice versa. In other words, a great deal of discussion as to L2 English pronunciation pedagogy holds true for pronunciation assessment, and vice versa.

In this chapter the term "pronunciation assessment" is used to refer mainly to "L2 English pronunciation assessment". In addition, given that the terms "comprehensibility" and "intelligibility" are used interchangeably in the literature to refer to almost the same concept, as explained in section 2.5.1.3 below, the term "intelligibility" is only used to refer to both terms to maintain consistency.

2.2 Overview of Pronunciation Assessment

Pronunciation assessment, either formally or informally undertaken, is argued to be one of the most common forms of human L2 language assessment. According to Harding (2012b), it is perhaps the oldest recorded form of L2 language assessment, dating back to biblical times (Isaacs, 2018). According to Spolsky (1995), as explained in the Book of Judges—the seventh book of the Hebrew Bible and the Christian Old Testament (Davis, 2017)—Gileadites used a single-item phoneme test of the oral production of the word "shibboleth" to determine whether a person is a member of their own tribe or from the warring Ephraimites. The purpose of this old ancient test is similar to the purpose of a recent test: the L2 language Analysis for the Determination of the regional or social Origin of asylum seekers (LADO). LADO is an analysis to determine whether an asylum seeker has the right to apply for asylum based on his/her speech productions (McNamara, 2012).

2.3 History of Pronunciation Assessment

2.3.1 The Long History of Pronunciation Assessment

L2 pronunciation has had a long history in L2 language teaching and assessment (Isaacs, 2018; Setter & Jenkins, 2005). It "stood at the very beginning of language teaching methodology as a principled, theoretically-founded discipline, originating with the late-nineteenth-century Reform Movement" (Seidlhofer, 2001, p. 56). According to Isaacs (2018), at the beginning of the 20th century, promoters of the Reform Movement rejected the Grammar Translation Method (e.g., Sweet, 1899, as cited in Isaacs, 2018), with its exclusive concentration on the written form of L2 language, translation quality and correct grammar (Richards & Rodgers, 2014).

Instead, the advocates of the Reform Movement considered phonetics fundamental to teaching L2 languages, and they considered phonetic transcriptions a replacement for a native speaking teacher as model for the correct L2 pronunciation of an L2 language (Isaacs, 2018). The

Reform Movement brought together phoneticians interested in L2 pronunciation teaching from a number of European countries. This has led to the development of an interest in L2 pronunciation teaching, including in L2 English teaching, where this interest has lasted into the second half of the twentieth century (Howatt, 2004, as cited in Setter & Jenkins, 2005). The Reform Movement has also led to the establishment of the *International Phonetic Association* and the *International Phonetic Alphabet* (IPA), where the latter includes an inventory of symbols representing the sounds of all documented languages (Setter & Jenkins, 2005).

A very early instance of L2 pronunciation assessment that considered the values of the Reform Movement is the Cambridge English Phonetics test that was required for the original Certificate of Proficiency in English (CPE) in 1913 (Weir, Vidaković & Galaczi, 2013, as cited in Isaacs, 2018). This test was a compulsory written test where test takers (usually L2 language teachers) were required to phonetically transcribe written texts into both careful and casual L2 pronunciations (ibid.). The test also included items asking test takers to determine the place and manner of articulation for some segments (ibid.). However, this test did not last in this format (Isaacs, 2018). According to Weir et al. (2013, as cited in Isaacs, 2018), in order to make the test more appealing to test takers to attract more registrations, the test was amended to avoid the part on phonetics in 1932.

Following the importance that the Reform Movement gave to aural/oral skills, several articles in this regard were written by American authors in *The Modern Language Journal* in the 1920s to 1940s (Isaacs, 2018). A prevalent perspective adopted by those authors is that "the oral test will always be […] the only real one for pronunciation" (Greenleaf, 1929, p. 534). Years later, the need for speakers able to speak and understand L2 languages, besides the military need for this during the Second World War, led to development of Kaulfers's oral fluency test (Kaulfers, 1944). In this test, examiners rated oral production using two 4-level scales, where the first scale is for assessing the information conveyed, and the second scale is for assessing oral production quality, ranging from "unintelligible or no response" to "readily intelligible" (ibid.).

The special importance that Reform Movement advocates gave to phonetics, particularly segmental features and articulatory phonetics, existed in the same state all through the Audiolingual era in the 1950s and early 1960s (Isaacs, 2018). This is shown clearly in Lado's (1961) influential and important book, *Language Testing* (1961), that includes several sections on assessing the perception and production of L2 segments, word stress, and intonation. According to Isaacs and

Trofimovich (2017), more than half a century from the time of its publication, Lado's (1961) book is still the most comprehensive practical guide to pronunciation assessment, covering a range of topics on pronunciation assessment such as giving tests and scoring. This book is today an essential guide for creating many L2 pronunciation tests (ibid.). Lado (1961) repeated the idea of Lundeberg (1929), where both expressed their feelings of worry about there not being enough unbiased pronunciation assessment, and they recognized practical challenges that threaten the effectiveness of pronunciation assessment. For example, when it is not possible to conduct a face-to-face oral L2 pronunciation test (e.g., because of limitation of time or shortage of financial resources), Lado (ibid.) put forward the idea of indirect assessment using certain written responses (e.g., multiple choice) as a feasible alternative. Following Lado's (1961) recommendations, Buck (1989) tried to find out the practicality of using indirect written pronunciation assessment in a Japanese context by comparing it to an oral assessment. He found out that there were low correlations between the scores in the written pronunciation assessment and the ratings given to the test takers' oral productions (ibid.). He further found that there was low internal consistency among the items in both the written and oral assessment (ibid.). Despite the existence of features with the validity of indirect pronunciation assessment such the ones found in Buck's (1989) study, written L2 pronunciation tests that assess aspects such as segmental discrimination and stress placement still continue to be used in major tests around the world such as the Japanese National Center Test for University Admissions (Watanabe, 2013).

2.3.2 The Neglect of Pronunciation Assessment

In spite of the great importance given to L2 pronunciation in Lado's (1961) book, the attention to L2 pronunciation in L2 language testing lasted for a short time. Soon after the appearance of his book, when teaching techniques on L2 pronunciation (e.g., decontextualized drills) were going against the general direction of research at that time, L2 pronunciation began to be neglected (Celce-Murcia, Brinton, & Goodwin, 1996). The Naturalistic Approach to teaching, that became known near the end of 1960s and at the beginning of the Communicative era, and that was in existence until the 1980s (Isaacs, 2018), placed less emphasis on L2 pronunciation teaching, considering it useless or even interfering with L2 acquisition and with attaining communicative competence (e.g., Krashen, 1981). Similarly, the Communicative approach advocates held that L2 pronunciation is similar to any other linguistic form (e.g., grammar), and thus it could be acquired gradually through exposure to comprehensible input alone, and that formal instruction had no role in this regard, although researchers decades later revealed findings challenging those claims (e.g.,

Saito, 2012). According to Isaacs (2009), based on the early principles of the communicative approach, comprehensible input was assumed to be sufficient for L2 pronunciation to develop.

As a result, L2 pronunciation was neglected for decades in the realm of applied linguistics as a whole and in L2 language testing in particular (Baker, 2013; Derwing & Munro, 2009b; Isaacs, 2013, 2018), although the assessment of L2 speaking proficiency has been of main interest to researchers in the field of applied linguistics following the first discussions of the Communicative approach (Kang & Ginther, 2018). L2 pronunciation was also given insufficient attention in many EFL classrooms (Derwing & Rossiter, 2002; Foote, Trofimovich, Collins, & Urzúa, 2016). This is well demonstrated by research written by L2 pronunciation advocates and L2 pronunciation material developers beginning in 1990, where they have been referring to the neglect of L2 pronunciation in L2 English teaching and learning (e.g., Elliott, 1997; Gilbert, 2010; Rogerson & Gilbert, 1990, as cited in Isaacs, 2018). The fact of not giving L2 pronunciation enough care and attention was still recognized and acknowledged among researchers of L2 pronunciation until very recently (e.g., Baker & Burri, 2016; Buckingham, 2015; Chan, 2016; Couper, 2017; Monfared & Khatib, 2018), but this criticism has been considerably missing in the scope of pronunciation assessment; only a few researchers have very recently criticized the neglect of pronunciation assessment in L2 language testing and research within the field of applied linguistics (e.g., Bøhn & Hansen, 2017; Isaacs & Harding, 2017; Isaacs & Trofimovich, 2017). According to Isaacs and Harding's (2017) research timeline, Buck's (1989) aforementioned article was the only research on pronunciation assessment since 1989 until the emergence of PhonePass, the first existing automated L2 speaking test, in 1999 (Bernstein, 1999). L2 pronunciation has received limited treatment in speaking assessment books (e.g., Luoma, 2004). Even in standard texts such as the *Cambridge Language Assessment* series, pronunciation assessment has received limited treatment; there is no dedicated book on assessing L2 pronunciation within that series (Isaacs & Harding, 2017), and there is still no book on that at the time of writing this chapter. Moreover, in the realm of English as a lingua franca (ELF) research, there is a relatively small amount of research within L2 language assessment in general, let alone in pronunciation assessment (McNamara & Roever, 2006).

The neglect of L2 pronunciation instruction, and necessarily pronunciation assessment, could be linked with another generally reported phenomenon within teacher cognition research. The findings in this area of research generally show that there is a dearth of confidence, comfort, and/or training and education amongst teachers and even teacher trainers, sometimes accompanied by scepticism about the significance and value of L2 pronunciation teaching, which

results in ignorance of L2 pronunciation in L2 language classes (and necessarily in L2 language assessment) (e.g., Baker, 2011, 2014; Breitkreutz, Derwing, & Rossiter, 2001; Burgess & Spencer, 2000; Burns, 2006; Buss, 2013; Couper, 2016; Derwing & Munro, 2009b; Foote, Holtby, & Derwing, 2012; Foote, Trofimovich, Collins, & Urzúa, 2016; Fraser, 2000; Henderson et al., 2012; Jenkins, 2007; Macdonald, 2002; Morin, 2007; Murphy, 1997). For instance, there are some studies that found that teachers do not know how, nor have the confidence, to teach suprasegmentals, especially stress and intonation, in spite of their recognition of their significance (Baker, 2011; Burgess & Spencer, 2000; Foote et al., 2012). This will necessarily hold true for assessing suprasegmentals. Also, according to Jenkins (2007), neglect of L2 pronunciation teaching (and necessarily pronunciation assessment) among English NNSs arises from a lack of confidence in their own L2 variety of English. Even if those speakers have confidence in their variety of English, they may still need training in L2 pronunciation pedagogy that involves the production of sounds, stress, rhythm and intonation and how to teach them to their students (ibid.). As a result, in addition to this training, they may need further training for assessing these constructs. As for English NSs, the general findings within teacher cognition research show that their neglect of L2 pronunciation teaching is usually a result of both the lack of confidence in L2 pronunciation teaching and inadequate education about L2 pronunciation pedagogy (e.g., Baker, 2011; Fraser, 2001; Macdonald, 2002). This is complicated by the fact that EFL teachers, including both NNSs and NSs, may still have the challenge of teaching L2 pronunciation to students with a wide linguistic diversity (and assess their L2 pronunciation), who usually come with different L2 pronunciation needs that require effective prioritization in the classroom (and in assessment; Schaetzel & Low, 2009).

According to Thomson (2018), one reason behind the teachers' discomfort and confidence as regards teaching L2 pronunciation (and necessarily pronunciation assessment) is that L2 pronunciation is the least understood L2 language skill. While many teachers may believe in the significance of L2 pronunciation to L2 language acquisition, a very little research has tried to determine the elements of good L2 pronunciation (ibid.). Lado's (1961) text (above-mentioned) is an early exception to this, followed by Munro and Derwing's (1995) seminal work "Foreign accent, comprehensibility, and intelligibility in the speech of second language learners" which clarified the related but somewhat independent measures of L2 pronunciation (i.e., foreign accent, comprehensibility, and intelligibility). The concepts of those measures or constructs have had effect on Munro and Derwing's further research for two decades and on the field of L2 pronunciation in general (e.g., Derwing & Munro, 1997; Derwing & Munro, 2009a, 2009b; Derwing, Munro, &

Thomson, 2008; Munro & Derwing, 1995a, 1995b; Munro & Derwing, 1999; Munro & Derwing, 2011; Munro & Derwing, 2015; Munro, Derwing, & Morton, 2006).

The neglect of L2 pronunciation could also be associated with the significant time limitations that EFL teachers encounter in class as a result of the requirement to attend to numerous L2 language skills dictated by a rigid curriculum, within which L2 pronunciation is mostly given limited treatment (Baker & Burri, 2016; Derwing & Munro, 2005). Some studies within teacher cognition research found that many L2 language teachers perceived that both L2 pronunciation instruction and providing learners with feedback on their L2 pronunciation are too time-consuming (e.g., Foote et al., 2011; Khaghainejad & Maleki, 2015; Warren, Elgort, & Crabbe, 2009). This in return could result in neglect of L2 pronunciation as well. For example, Foote et al. (2012) found that L2 language teachers do not give enough class time to L2 pronunciation; many of the teachers in their study stated they spend less than 5% of their class time on pronunciation instruction, and some other teachers stated their spending of as little as 1% of the class time on pronunciation instruction. However, the neglect of L2 pronunciation may further be linked with the lack of instruction and assessment instruments. According to Macdonald (2002), many teachers may not have access to suitable instruments for teaching and assessing L2 pronunciation, and thus this may cause neglect of L2 pronunciation teaching in the classroom and assessing L2 pronunciation (ibid.).

2.3.3 The Resurgence of Pronunciation Assessment

According to Derwing and Munro (2005) and echoed by Isaacs (2018), since the mid-1990s, L2 pronunciation underwent a revival of interest among the researchers of applied linguistics with a particular interest in Second language Acquisition (SLA); various articles featuring different aspects of L2 pronunciation have been published in leading SLA journals such as *Studies in Second language Acquisition* and *Language Learning* (e.g., Carey, Mannel, & Dunn, 2011; Derwing & Munro, 2009b; Harding, 2012b; Purpura, 2016). Such aspects included L2 intelligibility, accentedness (or foreign accent), fluency, L2 speaker background characteristics, and the linguistic features of their oral productions (Derwing & Munro, 2015). For example, some research looked at listeners' perceptions of accented speech (Lippi-Green, 2012; Munro, 2003). Some others investigated the relationships among intelligibility and accentedness (Derwing & Munro, 1997; Munro & Derwing, 1995; see definitions of these terms in section 2.5.1 below). Also, there is some survey research that explored whether and how teachers integrate L2 pronunciation in the

classroom (e.g., Burgess & Spencer, 2000; Breitkreutz, Derwing & Rossiter, 2001). There is also some research that has revealed that L2 pronunciation instruction can be useful for improving the speech of English learners (e.g. Couper, 2003, 2006; Derwing, Munro, & Wiebe, 1997). Also, some studies have defined key issues, such as the role of various L2 pronunciation features in communication (Derwing, Munro, & Wiebe, 1998; Hahn, 2004).

Also, since the mid-1990s, L2 English pronunciation has received attention as regards its methodology. For example, L2 English pronunciation experts have developed methods for integrating L2 English pronunciation teaching within a communicative framework (e.g., Morley, 1994; Wennerstrom, 2001). L2 English pronunciation experts have developed communicative activities where L2 English pronunciation plays a main role in conveying meaning, which in turn has developed a greater interest in the teaching of suprasegmental features (Setter & Jenkins, 2005). Also, some research has looked at teachers' needs in terms of L2 English pronunciation materials and curricula (e.g., Breitkreutz et al., 2001; Macdonald, 2002; Murphy, 2011). A small amount of research has examined teachers' practices in the classroom and teaching materials as regards L2 English pronunciation (e.g., Baker, 2014). One last example of the attention given to L2 English pronunciation methodology is the research that discussed how to enhance teachers' confidence in teaching L2 English pronunciation in the classroom using some particular methods, such as practical training in teaching L2 English pronunciation (e.g., Gilbert, 2010). However, the attention given to L2 English pronunciation was not reflected in L2 English teaching practices (Isaacs, 2009), let alone pronunciation assessment. For example, Levis (1999) analysed the instructional materials on intonation and found that the developers of these materials overlooked contemporary research on L2 English pronunciation. Still in more recent times, "instructional materials do not fit the bill in terms of providing authentic, context-rich activities that provide focused practice for the specific area of L2 pronunciation to be targeted, nor do they always draw on research evidence" (Isaacs, 2009, p. 4), in spite of "an explosion in the number of teaching resources" for L2 English pronunciation during recent years (Baker & Murphy, 2011, p. 37).

Following 2004, several signs have shown that L2 pronunciation has taken up a central position in applied linguistics research (Isaacs, 2018). Such signs included the appearance of L2 pronunciation-specific issues in journals, symposia dedicated to L2 pronunciation, and the organization of a conference dedicated only to L2 pronunciation in 2009 (*Pronunciation in Second Language Learning and Teaching*) (ibid.). These signs also included accumulated evidence from research demonstrating the effectiveness of L2 pronunciation instruction, and the launch of *The*

Journal of Second Language Pronunciation in 2015 (ibid.). Furthermore, according to Couper (2017), there has been an increasing number of studies within the area of L2 pronunciation since Derwing and Munro's (2005) report about the lack of research-based advice on L2 pronunciation instruction. Some of these studies demonstrated the need for explicit L2 pronunciation teaching at all stages (Zielinski & Yates, 2014, as cited in Couper, 2017).

However, instead of keeping pace with contemporary research on L2 English pronunciation and accordingly providing focused practice for L2 English pronunciation, no specifically instructional materials on L2 English pronunciation were developed. According to Foote et al. (2016), a great deal of L2 English pronunciation teaching is done merely ad hoc and as corrective feedback, usually in the form of recasts, which are usually not perceived by L2 English learners as feedback in response to their errors. Also, according to Foote et al. (ibid.), research based on classroom observation found that L2 English pronunciation teaching included only listen-and-repeat activities, except for some limited research that found only experienced instructors made use of a variety of activities, such as listening discrimination (Baker, 2014).

In the field of pronunciation assessment, there has been growing research work in the past decade in comparison to a dearth of studies in earlier decades, developed from the research work conducted within other fields such as SLA and sociolinguistics and influenced by the important role that L2 pronunciation has in automated L2 speaking scoring (Isaacs & Harding, 2017). For example, in the *Language Testing* journal, there were no articles discussing L2 pronunciation during the period beginning in 1984 and ending in 1988, in comparison to 4.45% of all the articles on L2 pronunciation during the period between 1998 and 2009 (Levis, 2015, as cited in Isaacs, 2018). Also, in the journal of *Language Assessment Quarterly*, there are no articles that took L2 pronunciation as a point of discussion in its issues since its beginning in 2004 until 2011; on the contrary, five articles discussing different aspects of L2 pronunciation (e.g., accentedness and intelligibility) have been published in the journal between 2012 and 2017.

According to Isaacs (2018), this resurgence of research in L2 English pronunciation, following its neglect subsequent to Lado's (1961) book, was caused by a change of focus and thought since the Audiolingual period. This shift of focus and thought has led to increased emphasis on specific suprasegmental features and growing awareness of the critical need to improve teachers' knowledge of L2 pronunciation (Celce-Murcia et al., 1996). Levis's (2005) characterization of the two "contradictory principles" in L2 pronunciation teaching can help explain this shift of focus

and thought as regards L2 pronunciation (p. 370): the nativeness principle and the intelligibility principle.

2.4 The Nativeness Principle

The nativeness principle posits that nativeness (or nativelike pronunciation) is both a primary goal of L2 pronunciation learning and a standard for pronunciation assessment (ibid.). It entails that the goal of L2 pronunciation teaching should be to assist learners in getting rid of their foreign accents in order to sound more nativelike. Accordingly, nativeness is most regularly measured by assessing listener perceptions of how a speaker is (foreign) accented or not (foreign) accented on a Likert-type scale (Derwing & Munro, 2015).

The nativeness principle has held a strong position in L2 pronunciation teaching and assessment for a long time (Levis, 2005). Traditionally, nativeness in L2 learning and teaching was perceived as a measure of success in L2 acquisition, and the standard for L2 language pedagogy and L2 research (Angelovska & Hahn, 2009; Cook, 1999; Deterding, 2010). As McKay (2012, p.10) mentions, "traditionally L2 pedagogy and research have been dominated by the assumption that the goal of bilingual users of English is to achieve native-like competence in English." This approach is built on the Chomskyian notion that suggests the native speaker is the ultimate expert on L2 language use (ibid.). In practice, SLA research has regularly considered L2 learners in relation to the native speaker; success and failure in L2 have repeatedly be determined with reference to native speaker competence (Bley-Vroman, 1983).

2.4.1 Empirical Research Supporting Nativeness

There are some empirical studies in the literature supporting the consideration of nativeness as a goal in L2 English teaching and assessment. For example, Smith and Bisazza's (1982) study supports the role of exposure to native speakers in promoting intelligibility. In their study, American, Indian and Japanese speakers of English recorded paragraphs and phrases in English. Next, these recordings were given to university students in English as a Foreign Language (EFL) contexts (Japan, Taiwan, and Thailand) and English as a Second Language (ESL) contexts (Hong Kong, India, and the Philippines). The students were then asked comprehension questions based on the recorded paragraphs and phrases. The findings of this study reveal that the American speaker was the easiest to understand, while the Indian speaker was the most difficult to understand (ibid).

Similar to Smith and Bisazza's (1982) study, Ortmeyer and Boyle (1985, as cited in Dimova, 2018) found that there is an advantage of American and British English varieties to intelligibility. In their study, they collected recordings of American, British and Chinese speakers, and then had 228 students at the Chinese University of Hong Kong listen to those recordings and then do dictation tasks (ibid). The findings of this study show that the students scored higher on the tasks when listening to the American and British speakers compared to their score for the Chinese speakers (ibid).

However, taking these two studies and other similar studies (e.g., Bent & Bradlow, 2003; Ekong, 1982; Major et al., 2002) into consideration, Dimova (2018) argued that such findings could be attributed to the usual practice in L2 English language learning where L2 English language learners tend to have higher exposure to NS varieties, especially American English, compared to other English varieties. Similarly, according to Flowerdew (1994), familiarity with accents tend to affect intelligibility, and thus L2 learners generally find it difficult to understand unfamiliar varieties. Thus, such findings could also be attributed to familiarity issues. In support of Flowerdew's (1994) argument, Tauroza and Luk (1997, as cited in Dimova, 2018) conducted an experiment on 63 Hong Kong school students, where the students listened to Hong Kong English and Received Pronunciation (RP). The findings of this experiment reveal that, besides Hong Kong English, RP was easily comprehensible to the students. Tauroza and Luk (ibid.) argued that this result could be attributed to the prevalence of RP in instruction in Hong Kong. They concluded that even non-local varieties are easily comprehensible if they are familiar accents to the listeners (ibid).

2.4.2 Criticism of Nativeness

Nativeness has regularly been criticized by L2 pronunciation experts and applied linguists for many reasons. For example, nativeness is impracticable to attain for most L2 learners even after years of L2 immersion, with only very few L2 advanced learners being able to sound nativelike (e.g., Abrahamsson & Hyltenstam, 2009; Jenkins, 2000; Loewen, 2015; Moyer, 1999; Munro & Derwing, 2011; Singleton, 2005). This can be attributed to many causes including the influence of L1 (Flege, 2003), age of acquisition of the L2 (Flege, Munro & MacKay, 1995), quality and quantity of input (Flege, Yeni-Komshian & Lin, 1999), and motivation, attitude to and aptitude for L2 acquisition (Moyer 1999). In line with this, Baker and Burri (2016) explained that subsequent to the beginning of puberty, it seems that L2 learners may struggle more with attempting to achieve nativelike L2

pronunciation in English; thus, it seems unreasonable to aim for nativeness in L2 pronunciation as it is almost impossible to achieve.

Another criticism against nativeness is the feature of deciding whom should be considered a NS (Cook, 1999), since there is variation within native English varieties in segmental and suprasegmental aspects of speech, especially in vowels (Lindemann, 2017). Furthermore, Jenkins (2000) points out that "a native-like accent is not necessary for intelligibility in ELF interaction" (p. 207). Also, according to Kennedy (2011), since effective communication could be considered the ultimate aim for teaching L2 pronunciation, and since communication is interactive and dependent on context, concentrating on intelligibility seems more suitable. In a like manner, Rugesæter (2012, as cited in Sannes, 2013) highlighted that L2 production that contains L2 pronunciation mistakes is still intelligible; thus, the attention should be aimed at what L2 speakers say rather than how they say it. Similarly, McKay (2012) stated:

> Traditionally L2 pedagogy and research have been dominated by the assumption that the goal of bilingual users of English is to achieve native-like competence in English. However, for those individuals who use English essentially as a language of wider communication alongside one or more other languages they speak, achieving native-like competence is often not necessary or desired. (p. 10)

Furthermore, in the past few decades, nativeness has been challenged by the English as a Lingua Franca approach (Jenkins, 2000; Seidlhofer, 2001). This approach seems to have had an effect on L2 assessment and instruction by making them less native-speaker focused (Graddol, 2006). Nativeness has also been challenged by the growing research promoting acceptance of other varieties of English besides American and British Englishes (Jenkins, 2007, 2012, 2014; Kirkpatrick, 2010a, 2015, as cited in Monfared & Khatib, 2018; Matsuda, 2017, as cited in Monfared & Khatib, 2018; McKay & Brown, 2016, as cited in Monfared & Khatib, 2018).

One further criticism against nativeness is related to aspects of identity: some scholars argued that L2 learners may wish to retain identifiable traces of their national or L1 identity when they speak (e.g., Gatbonton, Trofimovich, & Magid, 2005; Giles, 1979; Rindal, 2010; Walker, 2010). According to Baker and Burri (2016), a person's identity is intricately and closely associated with his/her L2 pronunciation. Thus, when teachers try to assist L2 learners to sound nativelike, this could create the feature of forming negative judgments about these learners' identities since this

could suggest the teachers are trying to prevent these learners from disclosing their identities (ibid.). Instead, teachers should focus on assisting L2 learners to communicate effectively with other speakers of English rather than helping them conceal their identities by imitating NSs (ibid.).

Also in relation to identity, there have been criticisms against the validity of some standardized L2 English tests since they do not consider the sociolinguistic identity of examinees whose linguistic norms are different from the linguistic norms in the tests (e.g., Chalhoub-Deville & Wigglesworth, 2005; Davidson, 1994, 2006; Lowenberg, 1993, 2002). From an ideological perspective, Spolsky (1993) argued that the constant use of a single norm in testing, commonly either American English or British English, indicates the perpetuation of American and British world-views and cultures. In this regard he argued that "the prodigious power of the tests has supported the imperialism of the competing normative varieties of English" (p. 88). As a consequence of such arguments, some scholars recommended assessments that reflect the global spread of English, which require the examinees to use their own varieties and linguistic resources for successful communication in universal contexts (e.g., Canagarajah, 2006; Jenkins, 2002).

In this PhD study, EFL teachers' attitudes to nativeness in pronunciation assessment are investigated. It is hoped that the study will help ascertain if nativeness still has a strong position in language assessment, given the recent shift towards the acknowledgement of the importance of intelligibility in global contexts over sounding like an idealised NS. This is investigated through EFL teachers' language attitudes which include their readiness for action as covered in section 4.2.1, Chapter 4. Accordingly, if EFL teachers are oriented towards nativeness, this is likely to be reflected in their teaching and assessment practices: they would base their teaching and assessment on nativeness. Thus, it is important to know if EFL teachers' attitudes are oriented to nativeness, despite the criticisms against this construct demonstrated above. If they are oriented to nativeness, this would necessitate training and education for these teachers to change their attitudes. Their attitudes would need to be oriented to intelligibility, instead, in order to better prepare their students for communication in English as an international lingua franca. The significance of intelligibility over nativeness is demonstrated further in section 2.5 below.

2.5 The Intelligibility Principle

As a result of criticisms against nativeness, some scholars suggest that L2 speakers should develop their own distinct varieties of speech but with the focus on intelligible output (e.g., Jenkins,

Cogo, & Dewey, 2011). This is related to the argument suggesting that what an L2 speaker needs is to be understood rather than sounding as if (s)he is an NS (e.g., Derwing & Munro, 2009b; Isaacs, 2013). Accordingly, some researchers within the area of L2 pronunciation have stressed the significance of inclusion of only L2 pronunciation features crucial to intelligibility in L2 pronunciation instruction (e.g., Derwing & Munro, 2005; Field, 2005; Jenkins, 2000; Levis, 2005). As Field (2005, p. 399) put it, "arguably the most pressing issue in L2 pronunciation research today is the quest to identify the factors that most contribute to speaker intelligibility." Thus, it may not be meaningful or reliable to talk about nativeness as a measure of L2 pronunciation; instead, as many have argued, there is a need for emphasis on intelligibility, which does not necessarily correlate with a native accent (Munro & Derwing, 1995). This position has led most applied linguists to adopt Levis's (2005) second contrasting principle, the intelligibility principle, as the most appropriate goal for L2 pronunciation teaching, and, by extension, pronunciation assessment (Harding, 2012b). This principle holds that the chief goal of L2 pronunciation learning is for learners to be understood by their interlocutors (Levis, 2005; see section 2.5.1.3 below for more definitions of intelligibility), and consequently intelligibility emerges as an appropriate assessment criterion (Harding, 2017).

2.5.1 Definition of Intelligibility and Relevant Concepts

When Munro and Derwing's (1995a) article was published over 20 years ago, featuring the concepts "intelligibility, comprehensibility, and foreign accent," these scholars may have not expected the dramatic effect those concepts would have had on research. According to Google Scholar, Munro and Derwing's (1995a) article along with the other two relevant articles, Munro and Derwing (1995b) and Derwing and Munro (1997), have been cited over 3000 times. Therefore, while teachers and researchers may use and interpret those concepts differently, when considered together, these concepts stem from those early studies. In this section, Munro and Derwing's original definition of intelligibility will be provided, followed by definitions by other scholars, including the definition of intelligibility from a phonological perspective (Jenkins, 2000). In order to differentiate between intelligibility and the other relevant concepts of foreign accent and comprehensibility, the definitions and measures of intelligibility along with these relevant concepts, are also provided below.

2.5.1.1 Foreign Accent

There is no general agreement on the definition of foreign accent; neither Munro and Derwing (1995a) nor Derwing and Munro (1997) offered a clear definition. Munro and Derwing (1995a) only highlighted what it is not, where they stated that it is considered "a bad thing […] subject to treatment, intervention, or even eradication in […] the same way as a language pathology" (p. 74). Munro and Derwing (1995b) reemphasized this by defining foreign accent as "non-pathological speech that differs in some noticeable respects from native speaker pronunciation norms" (p. 289).

Munro and Derwing's understanding of foreign accent is based on the way they measure this construct, where they perceive a listener as a judge of the accentedness of a speech rather than using an instrument to evaluate how phonetically distinct the speech is from a particular target (Thomson, 2018). Munro and Derwing used English NS raters to assess the foreign-accentedness and comprehensibility of 10 Mandarin NSs using 9-point scales to see if there are interrelationships between comprehensibility and foreign accent. The 9-point scale used for rating foreign accent ranged from 1 = no foreign accent to 9 = very strong foreign accent (Munro and Derwing, 1995a, p. 79, 1995b, p. 294), or from 1 = no accent to 9 = extremely strong accent (Derwing and Munro, 1997, p. 5). Many recent studies measured the strength of foreign accent in ways similar to Munro and Derwing's methods, where all of them made use of scalar ratings to measure foreign accent (e.g., Kennedy & Trofimovich, 2008; Trofimovich & Isaacs, 2012; Kraut & Wulff, 2013; O'Brien, 2014; Saito et al., 2015, 2016). For example, Kennedy and Trofimovich (2008) had participants rate utterances varied along two dimensions: real-world expectations (true vs. false utterances) and semantic meaningfulness (meaningful vs. meaningless utterances). The participants rated the utterances using a scale of nine points, where 1 = no non-native accent and 9 = strong non-native accent. Similarly, O'Brien (2014) used a nine points scale, where 1 = no accent and 9 = extremely strong accent, which the participants in his study used to rate extemporaneous speech samples. Similar to Kennedy and Trofimovich (2008) and O'Brien (2014), Saito et al. (2016) had participants rate extemporaneous speech samples using a nine points scale, where 1 = no accent and 9 = heavily accented. Saito et al. (2015) used a digital slider of 1000 points where the participants in their study used it to rate the accentedness of extemporaneous speech samples. Although Trofimovich and Isaacs (2012) used a nine points scale where the participants in their study used it to rate the accentedness of extemporaneous speech samples, Trofimovich and Isaacs (ibid.) used the scale in the opposite direction, where 1 = heavily accented, 9 = not accented at all. Finally, and different to

the studies mentioned above, Kraut and Wulff (2013) had participants rate an academic passage read aloud using a seven point scale on how strong a speaker's accent is. These researchers did not mention the terminology they used for the different points of the scale (e.g., 1 = no accent).

There is also other research that has provided definitions of foreign accent similar to Derwing and Munro's definition in terms of considering foreign accent to be a "perceptual phenomenon" observed by listeners (Thomson, 2018, p. 16). For instance, Kennedy and Trofimovich (2008) defined foreign accent as "how closely the pronunciation of an utterance approaches that of a native speaker" (p.461), a definition repeated by O'Brien (2014). Likewise, Isaacs and Thomson (2013) interpreted foreign accent as "how different the speaker sounds from a NS" (p. 141). In a similar vein, Jułkowska and Cebrian (2015, as cited in Thomson, 2018) defined (foreign) accentedness as "the listener's perception of how closely the pronunciation of an L2 speaker mirrors the pronunciation of a native speaker of [… her/his own] language" (p. 212). Slightly different from those definitions, Saito et al. (2015) interpreted foreign accent as "listeners' perceptions of the degree to which L2 speech is influenced by his/her native language and/or colored by other non-native features" (p. 8).

2.5.1.2 Comprehensibility

In their most cited article, Munro and Derwing (1995a) did not offer a clear definition of comprehensibility; however, from the measurements they used for comprehensibility, it can be deciphered that comprehensibility is something separate from intelligibility (Thomson, 2018). While intelligibility is measured by how much of the speaker's message is understood at word and utterance level (see section 2.5.1.3 below), Munro and Derwing (1995a) measured comprehensibility by the effort listeners devoted to understanding utterances. To support this position, they cited Varonis and Gass's (1982) description of comprehensibility where they referred to it as "ease of interpretation" (p. 127). In the same vein, Munro and Derwing (1995a) argued that "two foreign-accented utterances may both be fully understood (and therefore be perfectly intelligible), but one may require more processing time than another" (p. 91).

Munro and Derwing (1995b) again defined comprehensibility as the listener's effort taken to understand an utterance, defining it as "listeners' perceptions of difficulty in understanding particular utterances" (p. 291). Similarly, Derwing and Munro (1997) stated that it can be defined as "judgements on a rating scale of how difficult or easy an utterance is to understand" (p. 2). Such

definitions state clearly that comprehensibility involves effort taken by a listener to understand a speech; somewhat confusingly, In Munro et al. (2006), it is referred to as "subjective intelligibility" (see section 2.5.1.3 below).

Munro and Derwing (1995a, 1995b) and Derwing and Munro (1997) measured comprehensibility through having listeners judge the comprehensibility of spoken utterances using a scale of 9 points. They used the same scale but with slightly different wording of the descriptors, but they did not mention the reason behind the change of wording in the latter two studies. For example, in the scale of Munro and Derwing (1995a, p. 79), point "1" represents "extremely easy to understand" and point "9" stands for "impossible to understand". In Munro and Derwing's (1995b, p. 293) scale, point "1" is for "not difficult to understand at all" and point "9" for "very difficult to understand". Finally, in Derwing and Munro's (1997, p. 5) scale, point "1" represents "extremely easy to understand" and point "9" is for "extremely difficult or impossible to understand.

Most recent studies followed Munro and Derwing's (1995b) notion of comprehensibility where they perceived it as the effort listeners take to understand an utterance (Thomson, 2018). For example, Kennedy and Trofimovich (2008) defined comprehensibility as "listeners' perceptions of how easily they understand an utterance" (p. 461), a definition repeated exactly in Isaacs and Trofimovich (2012). Jułkowska and Cebrian (2015, as cited in Thomson, 2018) stated that "[c]omprehensibility is concerned with listeners' impressionistic observation of how easy it is to understand the L2 speech utterance" (p. 212). Derwing, Munro, and Thomson (2008) defined comprehensibility as "the ease or difficulty with which a listener understands L2 accented speech" (p. 360). Similarly, Isaacs and Thomson (2013, p. 8) defined it as "ease or difficulty in raters' understanding of L2 speech". Most of the studies citing Munro and Derwing followed them in the way they measured comprehensibility (e.g., Isaacs & Thomson, 2013; Isaacs & Trofimovich, 2012; Kennedy & Trofimovich, 2008; Kraut & Wulff, 2013; O'Brien, 2014; Saito, Trofimovich, & Isaacs, 2015, 2016; Trofimovich & Isaacs, 2012).

2.5.1.3 Intelligibility

Munro and Derwing (1995a, 1995b) defined intelligibility as "the extent to which a speaker's message is actually understood" (p. 76, p. 291). This definition was repeated in Derwing and Munro (1997). Accordingly, Munro and Derwing treated intelligibility as the ability to correctly recognise individual words within an utterance (1995a, 1995b; Derwing & Munro, 1997). Similar to

Munro and Derwing's definitions, Nelson (1982) defined intelligibility as "the apprehension of the message in the sense intended by the speaker" (p. 63). Munro et al. (2006) divided this term into two aspects: subjective and objective intelligibility. They defined objective intelligibility as "the extent to which a speaker's utterance is actually understood" (ibid., p. 112). Subjective intelligibility is the same as "comprehensibility," i.e., "listeners' estimation of difficulty in understanding the message" (ibid., p. 112).

Other studies that looked at intelligibility defined it in a way similar to Munro and Derwing's (1995a) definitions of intelligibility. For example, Field (2005) defined intelligibility as "the extent to which the acoustic-phonetic content of the message is recognizable by a listener" (p. 401). According to Kennedy (2009), intelligibility is "the extent to which a speaker's message is understood" (p. 132). Kennedy and Trofimovich (2008) again defined intelligibility in a similar way. Jułkowska and Cebrian (2015, as cited in Thomson, 2018) similarly defined it as "the degree to which a speaker's production is usually understood by a listener" (p. 212). According to Levis (2006, as cited in Isaacs, 2018), intelligibility means the general understandability of L2 speech.

Some other researchers defined intelligibility at the level of segment rather than utterance. For example, Gooauych, Saito, and Lyster (2016) described intelligibility as "whether the sound can be considered as [the target consonant] or other [English] consonant" (p. 119). Bundgaard-Nielsen et al. (2012) described a segment as intelligible if "it was identified as belonging to the intended English category" (p. 651). However, those definitions are still consistent with Munro and Derwing's in terms of assessing intelligibility through listeners' recognition of speakers' oral messages (Thomson, 2018).

Although intelligibility has commonly been dealt with as a facet of L2 speakers, some researchers have interpreted it as "hearer-based" (e.g., Fayer & Krasinski, 1987), where they have assumed that listeners are responsible for successful communication. Other researchers emphasized that intelligibility is a two-way process where interlocutors reach intelligibility through interaction (e.g., Field, 2005; Smith & Nelson, 1985). However, in practice, the perception of intelligibility as a two-way process is not represented in most pronunciation assessment research (Isaacs, 2018a). Accordingly, Moyer (2013) stated that the nature of communicative breakdowns, strategies for avoiding such breakdowns and successful communication (e.g., phonological accommodation: i.e., change of phonological features to express solidarity or social distance) have

not been investigated enough, except for Jenkins's (2000) work based on which she has drawn the LFC (see section 3.5, Chapter 3, for more details).

However, some researchers defined intelligibility as similar to Munro and Derwing's (1995a) concept of comprehensibility (see section 2.5.1.2 above). For example, Abercrombie (1949, p. 120, as cited in Dimova, 2018) stated that "language learners need no more than a comfortably intelligible pronunciation." He further argued that accurate L2 pronunciation is not required for everyone except for spies and teachers, but all that is needed for L2 learners is to have easily intelligible pronunciation (ibid., as cited in Dimova, 2018). Accordingly, he defined intelligible pronunciation as that is readily understood by listeners, with "little or no conscious effort" (ibid., p. 120, as cited in Thomson, 2018). Similarly, Saito et al. (2015) defined intelligibility as "listeners' subjective perception of how much or how easily they understand an L2 speech" (p. 440); they referred to ease of understanding instead of ability to understand in defining intelligibility. This confusion between intelligibility and comprehensibility is highlighted by Levis (2006, as cited in Thomson, 2018), who stated that intelligibility is frequently used interchangeably with comprehensibility (ibid.).

Phonological intelligibility has also been emphasised in ELF pronunciation research. For example, Jenkins (2000) gave significance to the phonological features of speech in intercultural communication in her intelligibility paradigm (i.e., the LFC, see 3.11, Chapter 3). She argued that, despite of the significance of lexicogrammatical and pragmatic meanings, L2 pronunciation (represented by the phonological features of speech) is the very first noticeable aspect of speech (ibid.).

Many studies measured intelligibility in the way Munro and Derwing measured it. For example, Jułkowska and Cebrian (2015, as cited in Thomson, 2018) and Kennedy and Trofimovich (2008) made use of transcription tasks. Bundgaard-Nielsen et al. (2012), Hayes-Harb et al. (2008) and Thomson (2011) used forced choice tasks, where listeners are required to identify the sounds they understand. However, Hahn (2004) measured intelligibility through listeners' ability to paraphrase the main idea of a speech and their ability to respond to comprehension questions about this speech. However, Hahn (ibid.) used the term "intelligibility" to point out constructs that are largely related to comprehensibility and listening comprehension. In reality, Hahn (ibid.) did not state clearly that she was operationalising intelligibility in the study tasks, nor did she define this

term. Only the title of the article and her discussion of the findings suggested she was measuring intelligibility.

Munro and Derwing (1995a) referred to four techniques adopted by researchers to measure intelligibility: (1) correct transcription of words in an utterance; (2) the amount of key words identified by a listener; (3) correct paraphrasing; (4) rating scales. Then, towards the end of their article, they (ibid.) put forward a suggestion for further research into more controlled methods for assessing L2 pronunciation intelligibility. As a result, Kang et al. (2018) undertook an exploratory study to compare five techniques for assessing L2 pronunciation intelligibility: (1) a cloze exercise where listeners hear nonsense sentences (e.g., "A clean soul tapes their keys.") and they are required to transcribe the content words of those sentences; (2) a cloze exercise where listeners hear and then transcribe parts of true/false sentences, with the neighbouring parts low-pass filtered in order that only prosodic cues are used as a context; (3) a true/false sentence verification task where listeners are briefly presented with simple sentences and asked to judge whether they are true of false; (4) scalar ratings where raters use a 100-point slider to determine the percentage of words they understand from true/false sentences; and (5) transcription task using nonsense sentences in which listeners transcribed the entire sentence. Kang et al. (ibid.) found that a cloze exercise with nonsense sentences was the best technique to predict the listener's ability to understand long, meaningful sentences produced by the same L2 speakers, while scalar judgements were the worst technique to predict that ability.

2.5.2 Empirical Research Supporting Intelligibility

Some empirical research has found results demonstrating the relevance of considering intelligibility as a goal in L2 pronunciation teaching and assessment instead of nativeness. For instance, Ballard and Winke (2017) investigated the interplay between speakers' accents and intelligibility and their acceptability as EFL teachers, focusing on non-native listeners. The findings show that non-native listeners did not seem to readily label accented speakers as unacceptable teachers; instead, they associated the acceptability of speakers with their intelligibility (ibid.). These findings echo previous work by Derwing and Munro (2009a), which showed a similar result for non-native English-speaking engineers in an English-medium workplace setting. Thus, the participants in these two studies seemed to consider intelligible speakers acceptable speakers, even if they were accented speakers.

Furthermore, some other empirical research indicated that a strong foreign accent does not necessarily hinder intelligibility (e.g., Derwing & Munro, 1997; Derwing & Munro, 2015; Munro & Derwing, 1995; Munro & Derwing, 1999). According to Gluszek and Dovidio (2010), the claim suggesting that a foreign accent hinders communication and lowers intelligibility is associated with many unconscious negative stereotypes. For example, some research has found that individuals who have foreign accents are perceived as less intelligent (e.g., Bradac, 1990; Lindemann, 2003; Rubin, Healy, Gardiner, Zath, & Moore, 1997), or less competent (e.g., Boyd, 2003; Bresnahan, Ohashi, Nebashi, Liu, & Shearman, 2002), and so on.

Smith and Rafiqzad's (1979) study is another example of the empirical research that supports intelligibility. In their study, Smith and Rafiqzad (ibid.) included 1386 participants from 11 different countries, and they found that NS phonology is not more intelligible than NNS phonology. Their findings suggest that recordings of some readers, such as American and Hong Kong Chinese readers, were least intelligible to the participants, whereas recordings of readers from India, Japan and Malaysia were the most intelligible (ibid.). Smith and Rafiqzad (ibid.) concluded that intelligibility does not seem to correlate with how much an utterance is foreign-accented.

However, despite this research evidence in support of the intelligibility principle, both teaching practices and most empirical L2 pronunciation research continue to be based on the nativeness principle (Thomson & Derwing, 2015). This obvious contradiction could be caused by the fact that L2 pronunciation research, especially pronunciation assessment, is still developing (Thomson, 2018). This in return led to misunderstanding of the meaning of the intelligibility principle among many teachers and researchers, and unawareness of the extent to which a foreign accent could or could not cause communicative breakdowns (ibid.). Another cause for this contradiction could be confusion among researchers and teachers in terms of the interpretation and use of the terms "accent, intelligibility and comprehensibility" (ibid.).

2.5.3 Linguistic Features Important for Intelligibility

The position that intelligibility should be the major priority in L2 pronunciation normally raises the question of which linguistic features should be considered in L2 pronunciation teaching and assessment. A main challenge to researchers in L2 pronunciation research is how to empirically identify the linguistic features important for intelligible speech so that those features can be emphasized in L2 pronunciation teaching and assessment (Isaacs, 2018).

There is much research that has tried to determine these features. For example, on a large scale, Derwing and Munro (1997) explored extensively a broad range of linguistic features (e.g., grammatical, lexical, semantic, and phonemic and phonological errors) that affect the NS listener's global perception of three different aspects of L2 pronunciation (i.e., accentedness, comprehensibility, and intelligibility) using transcription tasks. On a small scale, some other research explored what speech features (e.g., segmentals and suprasegmentals) have a great influence on the intelligibility of L2 pronunciation. However, this section will focus on speech features only since they are the focus of the current PhD study.

2.5.4 Speech Features Important for Intelligibility

Although there have been some attempts to decide on which speech features are essential for achieving intelligible pronunciation, these attempts did not produce conclusive results (Saito, 2014). Some research has demonstrated the importance of teaching both suprasegmentals and segmentals for L2 pronunciation pedagogy (e.g., Baker, 2011; Burgess & Spencer, 2000; Derwing & Munro, 2015; Derwing, Munro, & Wiebe, 1998; Foote, Holtby, & Derwing, 2012). For example, Derwing and Munro (2015) emphasized that "both segments and suprasegmentals have a place in the pronunciation classroom" (p. 9).

Following the Communicative era, L2 pronunciation teaching has given an increased focus on suprasegmentals (Celce-Murcia et al., 1996). Some researchers during this time suggested that suprasegmentals have a bigger impact on the intelligibility of L2 pronunciation than segmentals (e.g., Anderson-Hsieh, Johnson, & Koehler, 1992; Derwing, Munro, & Wiebe, 1998; Grant, 2014; Hahn, 2004). For example, Grant (2014) argued that suprasegmentals are essential features of L2 pronunciation besides other paralinguistic and global features including voice volume, fluency, and voice quality. Catford (1987) suggested that suprasegmentals including word stress, prominence and thought groups are important for the intelligibility of L2 pronunciation, but he also suggested that functional load (a segmental feature) is important.

Functional load is a theory associated with segmental-based instruction (Brown 1988). This theory can be used to help decide which feature contrasts, if any, should be the target of L2 pronunciation teaching and assessment (ibid.). It provides predictions about the communicative effect of mispronunciations of some English sound contrasts (ibid.). This theory hypothesizes that in order to determine the error gravity of minimal pairs, a series of key factors should be taken into

account, such as (a) the frequency of minimally paired words, (b) segmental positions within a word, and (c) the likelihood that the minimal pair contrast is supported in regional English dialects, since listeners are more likely to be able to make perceptual adjustments for contrasts that are distinct in different regional varieties than for those that are not (ibid.). As Saito (2014) put it:

> For example, whereas the contrast of /ɹ/-/l/ in word initial positions distinguishes relatively many words (e.g. 'lead' vs. 'read' 'lock' vs. 'rock'), the contrast in word final positions (e.g. 'wall' vs. 'war' 'tall' vs. 'tore') remains unclear according to the regional varieties of English (e.g. the rhotic /ɹ/ is typical of General American but not Received Pronunciation). Thus, the contrast in word initial positions has a high functional load. (p. 252)

Derwing and Munro (2015) also reported the effect of functional load on the intelligibility of L2 pronunciation. They stated that the substitution error that occurs by pronouncing /ɪ/ as /i/ (e.g., "chip" as "cheap") is more likely to cause a communication breakdown than the substitution error that occurs by pronouncing /θ/ as /f/ or /t/ (e.g., "three" as "free" or "tree"), since the contrast between /ɪ/ and /i/ has a high functional load while the contrast between /θ/ and both /f/ and /t/ is much lower by comparison.

Also, some empirical research has tried to determine the speech features essential for intelligibility, but this research considered NSs the sole judges of what is intelligible and what is unintelligible (e.g., Crowther et al., 2015; Hahn, 2004; Kang, 2010; Winters & O'Brien, 2013; Zielinski, 2008). In these studies, the researchers employed English NSs as raters of English NNS L2 pronunciation, and then based on the NS ratings the researchers determined different speech features as important for intelligibility.

For example, Crowther et al. (2015) found that only segments are essential for intelligibility. Crowther et al. (2015) investigated the recordings of a narrative task performed by three groups of English NNSs. Each group has a different L1 background and included 15 speakers. These NNSs were international undergraduate students at a Canadian university. The NNSs were rated by 10 English NSs, who were either graduate students or held an MA or a PhD. All raters were familiar with English accented speech and they had teaching and linguistic backgrounds. In a part of their study, Crowther et al. (ibid.) had the raters assess the recordings in terms of intelligibility and accentedness and in terms of pronunciation factors including segmentals, word stress, intonation

and rhythm. Later, Crowther et al. (ibid.) analysed the correlations between the ratings of intelligibility and accentedness and pronunciation factors. In a part of this investigation, Crowther et al. (ibid.) found that pronunciation intelligibility was significantly correlated with segmentals only, and only for one of the three NNS groups.

However, some other research found that primary stress—i.e., the most noticeable syllable in a word: it is a longer, louder and higher in pitch (Hahn, 2004)—is essential for intelligibility. For example, a part of Hahn's (2004) study examined NNSs' reactions to non-native primary stress in English in three different cases: primary stress correctly placed, incorrectly placed or completely missing. The stimuli for the study experiment included a recorded text with three versions different from each other according to the three cases of primary stress. Hahn (ibid.) recorded a NNS with background knowledge on phonetics reading the different versions of the text. Each version was presented to 30 NS undergraduate students (the total of the students was 90). These students were similar in terms of the amount of exposure they had had to non-native speech. Hahn (ibid.) measured overall comprehension of the recorded text in two ways: having the students write down as much as they recall from the recordings, and through a short comprehension quiz. The results show that the students recalled significantly more content when primary stress was correctly placed than when it was incorrectly placed or entirely missing.

Also, Zielinski (2008) found correlation between primary stress and intelligibility, but she also found that segments are essential for intelligibility. Zielinski (ibid.) investigated the features of speech the listeners rely on in identifying the intended words by speakers. She also looked at the non-standard features of speech that mislead the listeners into wrongly identifying the speakers' intended words. Zielinski (ibid.) had three English NSs listen and orthographically transcribe utterances taken from connected conversational speech produced by three NNSs. The NNSs came from different L1 backgrounds (Korean, Mandarin and Vietnamese). To determine the speech features the listeners depend on to identify the speakers' intended words, Zielinski (ibid.) firstly determined where the listeners experienced difficulty identifying the speakers' intended words. Then she determined whether the speech features used where the listeners experienced difficulty varied from the standard English phonology. The findings show that all three listeners relied heavily and consistently on the syllable stress pattern (the number and pattern of strong and weak syllables), regardless of whom speakers they were listening to. The findings also show that the listeners relied more consistently on segments in strong syllables than weak syllables. However, the findings show that non-standard syllable stress patterns and non-standard segments misled the

listeners into wrongly identifying the speakers' intended words, but non-standard segments in strong syllables were more misleading for the listeners than non-standard segments in weak syllables.

Finally, Winters and O'Brien (2013) found that suprasegmentals are relevant to intelligibility. In a part of their study, Winters and O'Brien (ibid.) investigated the impact of suprasegmentals on the intelligibility of English speech. They recorded natural speech produced by two English NSs, who read 24 sentences in English. Then, they manipulated the suprasegmentals in the sentences by transferring native intonation and syllable durations to non-native intonation and syllable durations. These manipulated stimuli were then given to English NSs to rate them in an intelligibility task. The English NSs were university students between the ages of 18 and 30. The results show that the more the suprasegmentals and segmentals were manipulated, the more the speech was perceived as accented, and the less intelligible the speakers became to the English NS raters.

However, the researchers of these studies considered intelligibility a one-way process, where NNSs are trying to make themselves intelligible to NSs who have the right to decide what is intelligible or not. More and more English NNSs are interacting in English with other NNSs, and in such cases English NNSs have the prerogative to decide what is intelligible or not (Smith, 1992; Smith & Nelson, 1985). Also, according to Lippi-Green (2012), considering English NSs as the sole judges of intelligibility promotes a standard language ideology that privileges English NSs whose language is considered standard in spite of the significant variations it has, while it neglects English NNSs whose language is not considered standard. According to Sewell (2017), interactions in English involve speakers from diverse L2 linguacultural backgrounds; even NSs of English are themselves extremely diverse as a result of migration and inequality, among other factors (ibid.). In this respect, and as early as the 1950s, Marckwardt (1958, as cited in Smith, 2015) suggested that NSs (not only NNSs) should modify their speech to communicate successfully as they speak different English varieties, and thus they may not understand one another. Accordingly, all NSs (not only NNSs) of English are, to a large extent, lingua franca users, even if they live in so-called monolingual environments (Sewell, 2017). As a result, there is a need for the adoption of an alternative approach that considers this diversity among English speakers (Jenkins & Leung, 2013). A promising solution is provided by a lingua franca approach to L2 English pronunciation teaching and assessment (Sewell, 2017). Since this PhD study adopts a lingua franca approach to L2 English pronunciation intelligibility, this approach is explained in detail in the following chapter (Chapter 3).

2.6 Pronunciation Rating Scales

This section is relevant to this PhD study for two reasons: (1) it is related to previous sections where it will extend the description of intelligibility by discussing its representation in L2 pronunciation rating scales; and (2) one of the issues this PhD study has investigated involved whether EFL teachers' attitudes were affected by assessment policies, which necessarily includes rating scales. Thus, it is ideal to give a background on rating scales.

Pronunciation assessment is usually integrated in speaking assessment, either in speaking rubric descriptors or in L2 pronunciation (or fluency) subscales within speaking rubrics (Dimova, 2018). The role of L2 pronunciation in speaking assessment has always been considered fundamental although its measurement tended to be "inconsistent and vague" (ibid., p. 55). While some scales do not have pronunciation references (e.g., the Common European Framework of Reference) (Council of Europe, 2011; North, 2000), other scales are "strikingly random in describing how pronunciation contributes to speaking proficiency" (Levis, 2006, p. 245, as cited in Dimova, 2018). For example, the rating scale in the Test of Spoken English, which is usually used for evaluating candidates for international teaching assistant (ITA) positions at US universities, has pronunciation as a feature of speaking proficiency, but pronunciation is still undeveloped and there are no links to intelligibility and comprehensibility (Educational Testing Service, 1995, as cited in Dimova, 2018). Furthermore, the L2 pronunciation descriptors in some of these scales, which are integrated in speaking rubrics, are inconsistent; thus, this results in underrepresentation of pronunciation (Isaacs & Trofimovich, 2011). Even the term "pronunciation" in speaking rubrics is itself inconsistent since it may refer to segmental features or suprasegmental features (ibid.).

2.6.1 Intelligibility in Pronunciation Rating Scales

The term "intelligibility" exists in the descriptors of the scales of many L2 oral proficiency tests (Dimova & Jensen, 2013, as cited in Dimova, 2018). For instance, the scoring rubric for speaking in the Test of English as a Foreign language (TOFEL iBT) evidently connects the terms "pronunciation" and "articulation" with the terms "intelligibility" and "listener effort" (ETS, 2014). Likewise, the descriptors in the IELTS speaking test associate the term "mispronunciation" with the terms "L2 accent" and "difficulties for the listener" on the one hand, and they associate the term "intelligibility" with the terms "L1 accent" and "understanding" on the other hand (IELTS, n.d.).

Other tests address intelligibility but without referring to it by its name. For example, the Oral English Proficiency Test (OEPT), which is an oral test for English speaking proficiency designed for candidates for international teaching assistant (ITA) positions at US Midwestern universities, has holistic scale descriptors addressing both listener requirements and speaker performance, i.e., addressing both intelligibility and comprehensibility (OEPP, n.d.). Another example is the ESL Placement Test (EPT), which is also designed for ITA candidates. However, unlike OEPT, this test includes only intelligibility in its scale descriptors. The first section of this test includes a three-minute interview, where the rater (NNS or NS) evaluates the interviewee's spontaneous utterance based on their ability to understand each word in their speech (Isaacs, 2008), which is similar to Munro and Derwing's (1995a, 1997) measure of intelligibility.

However, similar to pronunciation, intelligibility is seldom evaluated separately in rating scales, but it is integrated in the speaking rating scales besides other constructs (e.g., fluency) (Isaacs, 2008). This suggests that intelligibility is operationalised through raters' subjective perceptions, which means that comprehensibility is what is being used as a criterion in speaking scales rather than intelligibility (ibid.).

2.6.2 Evaluation of Pronunciation Rating Scales

An intuitive approach to the development of rating scales for L2 pronunciation has resulted in scales of low quality (Isaacs et al., 2018). For example, the ACTFL (2012) rating scale refers to behavioural indicators that are too general and thus unclear. One of the descriptors in this scale for the "novice low" level is "unintelligible pronunciation," and one of the descriptors for the "intermediate low" level is "strongly influenced by L1" (ibid.). Similarly, the IELTS (public version, n.d.) rating scale refers to behavioural indicators and linguistic features that are too general and not specific for each level. For example, the scale defines Band 4 as the level that describes the speaker who can use only a limited range of pronunciation features while having pronunciation that causes some difficulty to the listener (ibid.). Also, the scale defines Band 8 as the level that describes the speaker who can use a wide range of pronunciation features while being easy to understand, and L1 accent has a minimal effect on intelligibility (ibid.). However, the TOFIC (ETS, 2010) rating scale is less detailed and less general. For example, TOFIC (ibid.) defines level 8 that has scale score of 190-200 as the one that describes the speaker whose intonation and stress are highly intelligible at all times, whereas level 4 with a scale score of 80-100 is for the speaker who has difficulties with intonation and stress.

In addition to the examples above-mentioned, the TOEFL iBT pronunciation rating scale gives a little more detail as regards the linguistic features and descriptors across the levels of the rating scale, but similarly it still seems too holistic and thus difficult to use. For example, score 4 on the speaking rubric means that the speaker may have minor difficulties with pronunciation or intonation patterns, but that should not affect overall intelligibility (ETS, 2014). Another example is score 2 on the same rubric where it means that the speaker's speech is basically intelligible, but with effort needed from the listener's side due to unclear articulation and poor intonation. Also, other pronunciation rating scales, such as the CEFR Phonological control scale (Council of Europe, 2011) and the Cambridge ESOL common scale (not in use now; Taylor, 2011), are similar to those mentioned above in terms of being general and unclear, and in terms of associating foreign accent with unintelligibility.

2.6.3 Rating Scales Used in Pronunciation Research

Like the rating scales used in L2 pronunciation assessment, the rating scales used in L2 pronunciation research are also imperfect since their descriptors cannot be considered alone without comparison to a particular standard (e.g., very accented/not accented; very easy/difficult to understand) (Isaacs, 2018a). Although such scales have the advantage of suiting those listeners who have no previous linguistic training or rating experience, the validity of such scales is questioned since raters could interpret the scales' constructs differently since there is often absence of precise definition of those constructs (ibid.). For example, without providing a clear definition of comprehensibility, it could be interpreted by listeners as their understanding of every single word of a speech or their understanding of the general message (ibid.). Also, their attention could be turned to the amount of effort they believe they have devoted to understanding the utterance (i.e., "perceived cognitive load"), instead of establishing their ratings on the amount of information they have deciphering from the speech (ibid., p. 287). Therefore, in L2 pronunciation research that uses scales to operationalise pronunciation constructs such as intelligibility and comprehensibility, raters should be informed of the definition of such constructs and whether they should focus on word- or sentence-level understanding or on the overall message of the utterance (ibid.).

A notable example of L2 pronunciation research scales is Isaacs's et al. (2018) pronunciation scale for L2 English comprehensibility. It is a three-level scale that can guide raters on the linguistic features they need to consider at each L2 English comprehensibility level (Isaacs, 2018a). This scale was based on the Isaacs and Trofimovich's (2012) L2 English comprehensibility

scale guidelines. However, this scale is limited to speakers with a single L1 background and to a particular task type (i.e., picture narrative task) (Isaacs, 2018a). As a result of piloting this scale with focus groups of teachers, the scale was developed to become six-level scale suitable for the use on speakers with diverse L1 backgrounds preforming monologic academic speaking tasks (ibid.). The scale considered the problem of not having a standard with which a rater can compare a particular descriptor by providing examples of the errors that may hinder understanding at different proficiency levels (ibid.). A noteworthy feature of the scale is that sounding nativelike is not considered even at the highest level of the scale (ibid.). This was noted on the scale for the raters so that they are clarified that a detectable L1 accent in the speakers' L2 pronunciation does not prevent from reaching the top level of the scale (ibid.). However, this scale still requires further validation based on research data to make it line up with comprehensibility measures (ibid.).

2.7 Factors Affecting Pronunciation Assessment

A new area of investigation in pronunciation assessment is the factors that affect L2 pronunciation assessment, such as assessors' cognitive or affective variables (Mora & Darcy, 2017). For example, some research has found that L2 pronunciation assessors' knowledge of the assessed constructs, like L2 pronunciation fluency or accuracy, has a significant influence on their assessment (Browne & Fulcher, 2017; Harding, 2017). This could be the case for assessors who are L2 teachers as they would construct L2 pronunciation assessments based on their knowledge of the assessed constructs (Trofimovich & Isaacs, 2017). If they did not recognise what nuclear stress is, for example, they would not include it in their assessments.

There is also some evidence that familiarity with an L2 variety or a speaking task has an effect on L2 pronunciation assessment. In this vein, there are several studies that have investigated the effect of assessor's familiarity with an L2 variety or a speaking task on their behaviour using the speaker's overall score as the dependent variable (e.g., Carey, Mannell, & Dunn, 2011; Davies, Hamp-Lyons & Kemp, 2003; Huang et al., 2016; Isaacs & Thomson, 2013; Kim, 2009; Winke, Gass & Myford, 2013; Xi & Mollaun, 2011; Zhang & Elder, 2010). The results of these studies are contradictory; some studies show an impact of familiarity with an L2 variety on its assessment while others do not. Those findings parallel the findings from listening assessment research which have revealed that the familiarity with the speakers' accents in the recording listening tests give test takers an advantage, in comparison to the lack of familiarity with the speakers' accents (e.g.,

Harding, 2012a; Hsu, 2012; Major, Fitzmaurice, Bunta, & Balasubramanian, 2002; Ockey & French, 2016).

However, Isaacs (2018a) argued against the validation of studies that have not found a relationship between an assessor and the familiarity with the task or speaker's accent. In one such study by Huang et al. (2016), they did not find an effect of raters' accent familiarity on their L2 pronunciation assessment. According to Isaacs (2018a), plausible explanations for this result include statistical problems and the considerable overlap between the study groups as regards their exposure to the L1 of the speakers. This is in addition to Huang's et al. (2016) findings that suggest the raters perceived they were lenient in assessing the speakers since they were very familiar with their L1 accent (Isaacs, 2018a), which validates Harding's (2018) claim that familiarity with a speaker's accent may make an assessor more lenient when rating a speaker's performance (Harding, 2018). Thus, those points demonstrate the need for additional research work that avoids such methodological limitations (Isaacs, 2018a). In another study conducted by Xi and Mollaun (2011), no relationship was found between raters' accent familiarity and their L2 pronunciation assessment. According to Isaacs (2018a), factors such as the lengthy training that the raters received compared to other familiarity studies, and the very careful selection of the raters besides other precautions, may have interfered with the raters' scoring as those could have improved raters confidence and reduced familiarity bias (Isaacs, 2018a). Taking such arguments into consideration, it can be suggested, with some reservations, that L2 pronunciation tests should take rater familiarity into account when selecting raters and pre-test training (ibid.). Also, pronunciation assessment research should try to control for raters' accent familiarity (Winke et al., 2013).

As a result of the discussion as to the effect of listener's familiarity with the speaker's accent and/or the speaking task on pronunciation assessment, Isaacs and Thomson (2013) recommended the following: it is important to explain to the raters (after having background information on their experience) whether they should consider their own perspective when rating, or they should try to overlook their familiarity with the speakers' accent and/or the speaking task by, for example, acting as they are different target listeners.

In the context of WE, several studies have looked at the factors that have impact on the intelligibility of different English varieties (Dimova, 2018). Some of the factors that these studies have looked at included familiarity, accentedness, and attitudes to different varieties of English. Many of those studies compared varieties across Kachru's (1985) three circles of World Englishes

(Inner, Outer, and Expanding; see section 3.2.1, Chapter 3). As long ago as the middle of last century, Catford (1950, as cited in Dimova, 2018) recommended that there should be a threshold of intelligibility that indicates how much exposure to a variety is needed in order that the user of such variety becomes familiar with it, and thus s/he becomes more intelligible and has a more positive attitude towards it. To acquire better understanding of the relationship between familiarity and intelligibility, Gass and Varonis (1984) deconstruct the familiarity concept into four elements, where each of these contribute towards NS comprehension of NNS varieties. These four elements include: familiarity with the topic, familiarity with NNS pronunciation on the whole, familiarity with a specific NNS variety, and familiarity with a NNS speaker in particular (ibid.). They (ibid.) then tested this hypothesis and found that while the most effective element of the four on intelligibility was the familiarity with the discussion topic, all of these elements contributed towards intelligibility.

Another effect on pronunciation assessment is attitudes towards some accents. According to Yan and Ginther (2018, p. 78), "[a]ccent can elicit attributions that may bias the listener negatively against the speaker" (see Table 1.2 below for research samples on the effect of attitudes towards accents on pronunciation assessment).

Table 1.2: Samples of studies looking at the effect of attitudes towards accents on pronunciation assessment

Study	Type of effect
Cargile & Giles (1997); Giles (1972); Giles et al. (1995); Lambert (1967)	The relationship between the attitudes to accent from different angles.
Davies et al. (2003)	Teachers' acceptance of an L2 variety of English may have influence on their rating of L2 oral performances.
Dávila et al. (1993); De Klerk & Bosch (1995); Kalin & Rayko (1978, as cited in Yan and Ginther, 2018)	Perceptions of accents could affect assessment of pronunciation in terms of aspects such as power and socioeconomic status.
Hsu (2012)	Teachers' preference for a particular English variety over another one could affect their L2 pronunciation intelligibility judgment.
Kang (2012)	Raters' characteristics and prosodic features of speaker accentedness affect raters' judgements on L2 oral performances.
Schinke-Llano (1983, 1986, as cited in Yan and Ginther, 2018)	Existence of a foreign accent could cause underestimation of the L2 oral proficiency.
Thomson (2018)	Negative oral evaluation can stem from stereotypes connected with particular accents.
Trofimovich & Isaacs (2017)	L2 teachers who favour nativeness would prioritise pronunciation assessment founded on native speaker norms.

However, the attitude towards a particular accent could change over time (Yan & Ginther, 2018). In this vein, there is a hypothesis that entails attitudes towards a particular accent are subject to change because of gaining familiarity with that accent (e.g., Gass & Varonis, 1984; Sato, 1998). However, according to Yan and Ginther (2018), pronunciation assessment research has not reached general agreement on whether increased familiarity with a particular accent could result in higher intelligibility scores in rating and more positive attitudes towards that accent. For instance, Fayer and Krasinski (1987) investigated the attitudes of NS English and NNS (Spanish) English listeners to Puerto Rican learners of English with different proficiency levels. The findings of this study reveal that the NNS English listeners were less tolerant than the NS English listeners despite

the NNS English listeners shared the same L1 with the learners. Thus, according to Yan and Ginther (2018), current research may suggest that the attitudes to a particular accent could change over time due to familiarity with L2 accent but not in all contexts.

Another factor that could influence pronunciation assessment is listener L2 language proficiency and/or linguistic awareness. According to Yan and Ginther (2018), although this factor is important to the perception of (intelligibility of) L2 accented speech, it has often been neglected in research. Research on intelligibility rating tends to employ NS English listeners or very proficient English listeners for intelligibility judgement tasks (ibid.). However, intelligibility judgement or rating could be negatively affected if listeners have a low level of L2 language proficiency (ibid.). According to Yan and Ginther (ibid), the effect of listeners' L2 language proficiency and linguistic awareness should be approached together with the cognitive demand made to process the meaning of an utterance. In other words, the higher language proficiency and linguistic awareness the listeners have, the less cognitive demand they made to process the meaning of an utterance (ibid.).

Despite being indirectly relevant, Thompson's (1991) study shows that having extensive metalinguistic knowledge or awareness had an impact on the perception of L2 accent. In his study, Thompson (1991) employed two groups of English NSs (whom he classified as experienced and inexperienced NSs). The participants were asked to rate samples of speech performed by 36 Russian speakers of English with different levels of accentedness. The samples of speech were recorded by having the speakers read written pieces. The findings reveal that the words that have difficult sounds were perceived by the participants as more accented compared to the other normal speech. The findings also show that the experienced participants were more tolerant with accented speech of the speakers. According to Yan and Ginther (2018, p. 79), these findings make sense since the perception of intelligibility and accentedness coincide with meaning process, which both require "a certain level of cognitive processing load to the listener." They continued, "[t]herefore, when additional effort is required to decode the phonological, syntactic, or lexical information, the adjustment to L2 accent is likely to be laboured and the tolerance of the accent lowered" (ibid., p. 79).

One way of avoiding rater effects in pronunciation assessment is by considering the use of automated scoring systems (Isaacs, 2018a). However, in addition to the weaknesses of automated scoring mentioned above, those systems are not "foolproof" and susceptible to "error in the form of false positives (i.e., system scores the production of a correctly pronounced L2 sound as an error)

and false negatives (i.e., system fails to detect the production of an incorrect L2 sound […])" (ibid., p. 288). But, in order to reduce the effect of such shortcomings, future automated scoring systems and human raters could perhaps complement each other in one integrated scoring system (Isaacs, 2018b). For example, one approach to creating such a system is having the automated scoring system operationalise the features that it can measure most effectively (e.g., durational measures), while having human raters score the features that the automated scoring system is less capable of measuring (e.g., cohesion) (Isaacs, 2018a). This could tackle the limitation of automated scoring systems in being able to detect only specific linguistic features (ibid.). However, the linguistic features that the raters should focus on and how human scoring can complement automated scoring are still required careful consideration (ibid.). The use of both automated and human scoring has proved successful in measuring the skill of writing in TOEFL (Ramineni et al., 2012), thus the use of such approach could be observed soon in large-scale L2 speaking tests (Isaacs, 2018a), which inevitable include L2 pronunciation assessment.

2.8 Summary of Chapter

Since this PhD study looks at attitudes in pronunciation assessment, this chapter has introduced a detailed description of pronunciation assessment. It has firstly provided an overview on pronunciation assessment, and then it has addressed the neglect of pronunciation assessment. This line of discussion has demonstrated the need for further research into this neglected area, which this PhD study tries in part to fulfil. The chapter has then introduced discussions regarding two main contradictory principles in L2 English pronunciation teaching: nativeness and intelligibility. The chapter has demonstrated the significance of intelligibility over nativeness and the recent shift of attention to intelligibility. This PhD study investigates whether EFL teachers' attitudes are aligned with the literature in terms of the significance of intelligibility over nativeness in an assessment context. This is investigated through EFL teachers' language attitudes which involve their readiness for action as mentioned in section 4.2.1, Chapter 4. Accordingly, if EFL teachers are oriented towards intelligibility or nativeness, this is likely to be reflected in their teaching and assessment practice: they would base their teaching and assessment on intelligibility or nativeness according to their attitudes to both constructs. If the EFL teachers in this PhD study are oriented to intelligibility, this will demonstrate that the teachers commit themselves to the significance of intelligibility over nativeness and the shift of attention to intelligibility demonstrated in L2 research. If the teachers, otherwise, are oriented to nativeness, this will necessitate training and education for these teachers to change their attitudes to align with recent L2 research, given the

importance of readying L2 learners for communication in English as a lingua franca in global settings. The chapter has finally discussed pronunciation assessment scales and factors affecting pronunciation assessment. This line of discussion serves as a background that may help to understand the investigation of the effect of pronunciation assessment policies and other factors on EFL teachers' attitudes in this PhD study.

CHAPTER 3: LITERATURE REVIEW 2: LINGUA FRANCA APPROACH TO INTELLIGIBILITY

3.1 Introduction

Since L2 English pronunciation intelligibility in this PhD study is approached from the viewpoint of ELF, and it is measured using the LFC as a premise and a framework, this chapter presents a review of the literature on ELF, including the literature on the LFC. Firstly, the context of ELF is set out by providing an overview on this concept: how it has come into existence and how it indicates a change of the ownership of English. Then, language assessment and L2 English pronunciation in the context of ELF are addressed. Next, the chapter addresses the logic behind considering a lingua franca approach to L2 English intelligibility. It then addresses a major development of this approach in the field of L2 pronunciation: the LFC. The LFC is presented followed by a discussion of the empirical research that has replicated the LFC or has investigated it. The chapter also presents the arguments that have arisen in the literature and that have supported or rejected the LFC, along with the reasoning of all arguments. This is followed by a related discussion about the LFC: should the LFC be considered a model of pronunciation? Next, the chapter introduces phonological features suggested by some research as essential for L2 English intelligibility, noting that English NS raters were used as judges of L2 English intelligibility. Finally, the chapter concludes with a discussion on whether ELF can be considered a new variety of English.

3.2 Overview of English as a Lingua Franca

3.2.1 Globalization of English

The phenomenon of globalization has had a great influence on the spread of English in the world (Cogo & Dewey, 2012). This has contributed to the English language having a position in the world that has not been achieved by any other language (Frath, 2010, as cited in Zoghbor, 2018). It has become the official language in many countries, and it has been used as a lingua franca in many educational, technological and business contexts around the world (Crystal, 2003; Graddol, 1997, 2006). In return, contemporary advances in technology and demography have increased the ongoing internationalization of English language, and they have also changed the way English language is used and conceptualized (Khatib & Monfared, 2017).

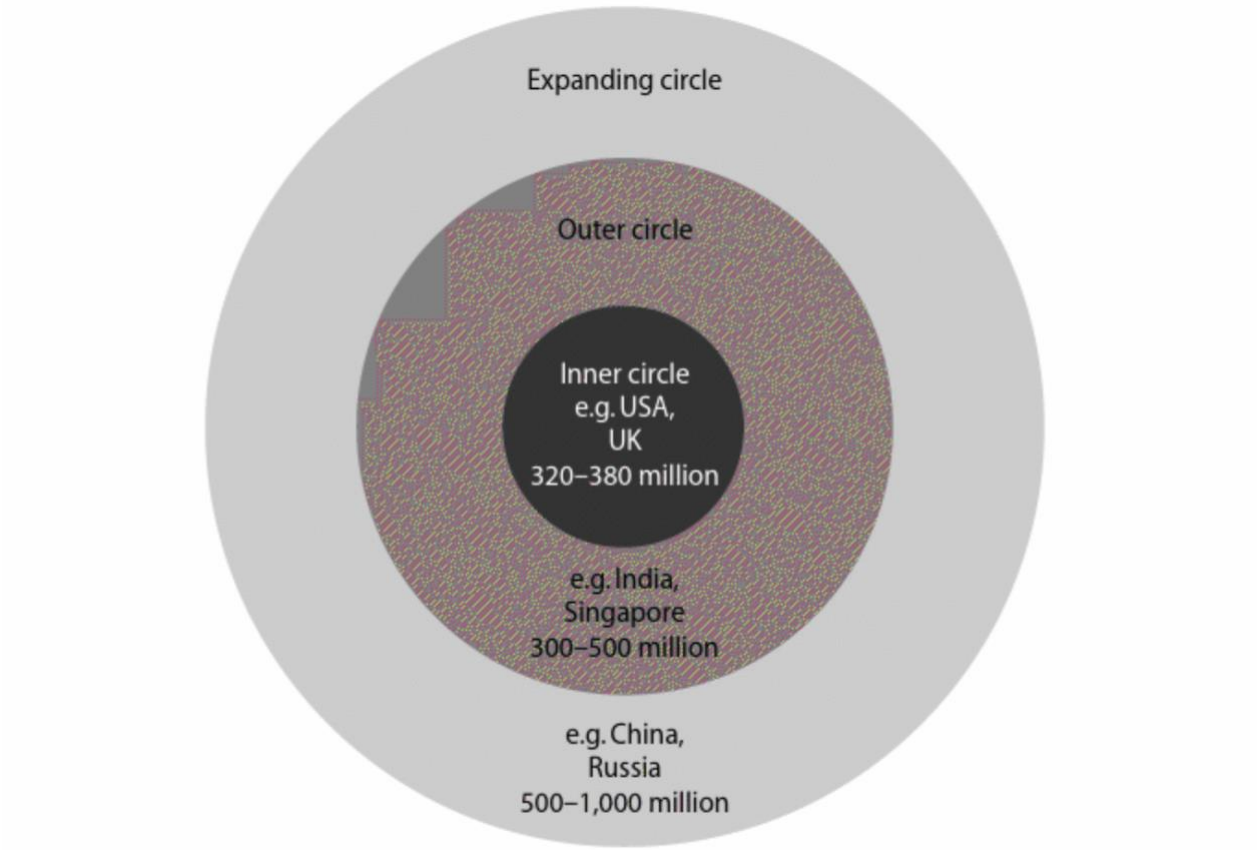
Unlike other languages, the English language has changed dramatically since about 1990 as it interacts more with other languages (Khatib & Monfared, 2017), and it is increasingly spoken by users from a wide diversity of backgrounds (ibid.). Kachru (1985) categorised these backgrounds into three circles according to the functions of English: the Inner Circle, the Outer Circle and the Expanding Circle (see Figure 1 for demonstrating of these circles). The Inner Circle includes the communities where English is spoken as a native language (e.g., UK, US, Australia, Canada, and New Zealand). The Outer Circle includes the communities which have a colonial history and use English as a second language (e.g., India and Singapore). The Expanding Circle represents the rest of the world, where English is spoken as a foreign language (e.g., Japan and Saudi Arabia). According to Beare (2015), one billion of the users of English are learners of English; 750 million of these learners come from the Expanding Circle, and a further 375 million of those learners come from the Outer Circle (ibid.).

However, the classification of countries into these categories seems only to help manage discussions regarding the spread of English worldwide; it does not strictly place countries in the right categories (Zoghbor, 2018). For example, in the United Arab Emirates (UAE), English is taking over the official language (Arabic), and it is used as a lingua franca (Boyle, 2011; Randall & Samimi, 2010). Generally, there is increasingly dependence on English in the countries of the Gulf Cooperation Council (GCC) caused by the demographical distribution of the populations in those countries (Zoghbor, 2014). For example, in the UAE there are more than 200 nationalities representing more than 80% of the total population in comparison to less than 20% of the locals (Al-Khouri, 2012). Those expatriates, who constitute the majority of the workforce in the UAE, come from non-Arab countries such as India, Pakistan, Bangladesh and the Philippines (Zoghbor, 2014; Boyle, 2011). The communication between these expatriates is established through the English language as a lingua franca and includes a great deal of variety in English. Accordingly, Graddol (1997) considered the UAE, besides other countries such as Lebanon, Somalia and Sudan, to be in the transition stage in terms of English language use, where the people of these countries are becoming second-language-users rather than foreign-language-users, and so do not fit neatly into Kachru's (1985) paradigm.

Overall, the classification of countries according to the type of English used in these countries seems more complex than the Kachru's (1985) "concentric model" suggests. The demographical distribution of the populations is changing rapidly due to the accelerated globalisation around the world which has led to presence of different speakers of English within a

single community. In line with this, Canagarajah (2006) stated there are more and more speakers of Outer Circle and Expanding Circle Englishes living in the Inner Circle countries; thus, World Englishes should not be viewed from the perspective of Kachru's (1985) "concentric model." Thus, the Kachru's (1985) "concentric model" may only be used to describe the English varieties existing within countries rather than the English varieties of the countries as a whole.

Figure 1: Kachru's three circles model for categorising English according to its functions, including the number of speakers in each circle (adopted from Crystal, 2012, p. 61)



3.2.2 Development of English into a Lingua Franca

As a result of the unprecedented predominant status that English language enjoys in this century as a global language, it has developed into a lingua franca in the realms of research, culture, business and economics (McKay & Bokhorst-Heng, 2008). According to Khatib and Monfared (2017), the spread of English worldwide, that has been caused by many distinct historical, political, economic and scientific factors, has developed new concepts such as "the acceptance of varieties of English, … cultural appropriateness of pedagogical materials, language and identity, ownership of

English, etc." (p. 214). It has also created multiple varieties of English, which has in turn given new names to English such as *English as an International English* (EIL; e.g., McKay, 2002, 2012), *English as a lingua franca* (ELF), often used interchangeably with EIL (Monfared & Khatib, 2018), *English as a global language*, *English as a world language* and *English as a medium of intercultural communication* (e.g., Crystal, 2008; Jenkins, 2000; Phan Le Ha, 2008; Seidlhofer, 2004, 2011) and *World Englishes* (WE; e.g., Kachru & Nelson, 2006; Kirkpatrick, 2007, 2010a, 2011).

Sharifian (2009, p. 2) defined ELF (or EIL, as both are often used interchangeably as mentioned above) as a paradigm shift in TESOL and SLA, which has happened as a response to the complexities connected to the extremely fast spread of English language around the world. Sharifian wrote:

> As a paradigm, EIL calls for a critical revisiting of the notions, analytical tools, approaches and methodologies within the established disciplines such as the sociolinguistics of English and TESOL, which explored various aspects of the English language. One of the central themes of EIL as a paradigm is its recognition of World Englishes, regardless of which circles they belong to. (ibid., p. 2)

In the field of ELF, the major focus is on communication instead of on the speakers' backgrounds, such as being a native or non-native speaker (NS or NNS, respectively) (Monfared & Safarzadeh, 2014). According to Kachru (1992), one of the main interests of ELF is acceptance of World Englishes without turning attention to the circle they are related to.

As for WE, it is used as an "umbrella term" to refer to two kinds of research (Hsu, 2012, p. 9). The first one is the "legitimacy, norms and usage of the new varieties of English", particularly referring to "nativized and institutionalized varieties in the outer circle" of Kachru's (1985) "concentric model" (ibid., p. 9). The second one is ELF, which has produced extensive research on the English language produced by the speakers of the expanding circle; it showed that the English language produced by these speakers seems not to be "random, irregular forms of English". As claimed by Seidholfer (2004, p. 212), ELF has "taken on a life of its own, independent to a considerable degree of the norms established by its native speakers, and that warrants recognition". However, there is disagreement among researchers regarding whether ELF can be considered as a part of World Englishes paradigm (see Berns 2008; Jenkins 2006).

3.2.3 Ownership of English Language

Within the area of applied linguistics, the ownership of English has recently been questioned and reconsidered (e.g., Brumfit, 2001; Crystal, 2000; McKay, 2006; Widdowson, 1994). For example, Widdowson (1994) argued that the English language no longer belongs to NSs, but to everyone who speaks it. Similar to Widdowson, Brumfit (2001) stated that English is used mostly by speakers whose first languages are different, and thus this has led to questioning the difference between the NSs and NNSs regarding who owns the English language (ibid.). Similarly, Matsuda (2003) argued that interaction in English commonly occurs among English NNSs, and thus this suggests a change in ownership of English language (where it no longer belongs to NSs). Also, according to Crystal (1997) and Graddol (1997, 2006), the use of English has extended beyond the interaction between NSs and NNSs, where the great majority of interactions using English do not involve English NSs. Statistically, interactions using English occur among more than one billion proficient speakers from Outer and Expanding circles, and thus this calls for reconsideration of the conventional assumption that English language belongs to its NSs (Beare, 2015).

3.3 English as a Lingua Franca and Language Assessment

The field of language testing and assessment has not carefully considered ELF in its research (Canagarajah, 2006a; Davidson, 2006; Jenkins, 2006a; Lowenberg, 2002). It has heavily relied on NS norms and has not considered English varieties spoken in many contexts (Davies et al., 2003), in spite of non-success in adopting NS norms in the international contexts (Lowenberg, 2002). This is the case up to the time of doing this PhD study (as far as the author of this study is aware). This seems to be the case even in the major international language tests. According to Newbold (2017), ELF is not reflected in the major international language tests used to assess university students before their access to higher education. Even standardized scales such as the Common European Framework for Reference (CEFR, henceforth; Council of Europe, 2001) still do not give ELF careful attention. According to Ahvenainen (2005), although the CEFR recognised the different levels of language competence, it still stresses the need for achieving native-speaker competence in the assessment criteria. A current example of this is as follows (retrieved from Council of Europe on 5th February, 2022): “Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party” (level B2).

Taylor (2005) highlighted the need for change in the assessment of English language in response to the growing number of English varieties, particularly NNS varieties. In this regard, she proposed that language testers should bear the changing status of English in mind when conducting assessments (ibid.).

As a result, two positions as regards the norm that should be used in L2 English testing have been adopted. These two positions represented different ideologies and approaches. The first one is the standard English position (e.g., Elder & Davies, 2006; Elder & Harding, 2008), and the second one is the ELF position (e.g., Canagarajah; 2006b; Davidson, 1993, 2006; Jenkins, 2002, 2006b; Lowenberg, 2002; Spolsky, 1993). The standard English position suggests that a single norm needs to be used to rate English language proficiency, while the ELF position is that considering a single norm to judge examinees on their English language proficiency ignores the linguistic richness in the different English norms used around the world and constitutes prejudice against examinees who use them (Hsu, 2012).

In support of the standard English position, Elder and Davies (2006) argued that certain requirements that guarantee test quality impose constraints on test development, and this affects the position towards adopting an ELF approach to designing language tests. These requirements include: (a) construct validity—i.e., what to measure given the test purpose and context; (b) fairness—i.e., providing bias-free scores regardless of the examinees; and (c) acceptability to stakeholders and examinees (Elder & Harding, 2008). For construct validity, Elder and Harding (ibid.) argued that the uncertainty regarding the appropriate language use and the standards used to assess the speaking performance makes testing agencies depend on standard English varieties for a higher degree of certainty. Similar to this is the uncertainty regarding the examiners' ability to attend to multiple varieties of English at the same time during oral proficiency tests (Davies et al., 2003). As regards fairness, Elder and Harding (2008) argued that standard English varieties are the varieties that most examinees are familiar with, but the examinees are different in terms of their familiarity with other varieties; thus, there would be a bias against the examinees who are not familiar with certain varieties, in contrast to the examinees who are familiar with these varieties (ibid.). Therefore, the standard English varieties are a safe and neutral approach (ibid.). As for acceptability to stakeholders and examinees, Elder and Harding (ibid.) argued that the strong ideological position taken by stakeholders and examinees in favour of standard English varieties justify their use in testing.

It is, however, fairly straightforward to counter Elder and Harding's position. For example, regarding construct validity, this can be avoided by applying one paradigm that combines all the varieties of English. An already established paradigm that combined NS and NNS English varieties is the Lingua Franca Core (LFC; see section 3.5 below). Since the LFC is an innovative paradigm and it is almost the only paradigm of its kind, it has received some negative criticisms (see section 3.5.3 below for these criticisms); however, such a paradigm can still be developed to achieve construct validity. As regards fairness, exposing examinees to standard varieties only may help the examinees pass an exam based on standard varieties, but it may not help them communicate successfully with English NNSs. Thus, to achieve real fairness, firstly examinees need to be exposed to intelligible NS and NNS varieties so that they become familiar with these varieties, then the exams need to be constructed based on these varieties.

ELF researchers have argued strongly against the standard English position (Dimova, 2018). They have argued that the standard English position stands against the local norms and perceives them as impaired varieties (ibid.). In contrast, the ELF position "allows for the realistic representation of different varieties including English varieties from the *Outer Circle*, i.e., post-colonial countries" (ibid., p. 51). This latter position validates Outer Circle varieties and strengthens their position to the extent of "their codification" (ibid., p. 51). However, while some ELF researchers argued for codification and standardisation of Outer Circle varieties, this group of ELF researchers did not recommend that for Expanding Circle varieties, which they perceived as varieties used for communication with native speakers only (Bolton, 2004). However, taking into consideration the dominant NNS to NNS communication using English, recently some other ELF scholars argued that the traditional approach to ELT, that observed the Expanding Circle varieties as dependent norms that should rely on standard English for communication, is no longer acceptable (Dimova, 2018). As a result, these ELF scholars argued that an ELF approach which considers communication among all users of English regardless of their varieties should replace the traditional approach to ELT (ibid.).

In language testing, ELF researchers have argued that the incorporation of varieties other than the "standard" ones in standardised tests could reduce the bias against NNS test takers and have a positive effect on language teaching and learning as a whole (Dimova, 2018). Also, some of the scholars who adopted the ELF position argued from a critical ideological point of view that imposing only a single norm for rating examinees could promote US or UK imperialism. For example, Spolsky (1993) argued that English tests have long been used to support American and

British Global-views and cultures. He considered this to be imperialism which ignores the linguistic richness present in the ELF speaker groups and supports particular norms that do not represent the diverse English used around the globe. Also, Davidson (1993, pp. 119-20) argued against what he called "the prevalent imperialism of major international tests of English" and that these tests are made by testing agencies that are so powerful that it is difficult to produce ELF-oriented tests in place of them. He said:

> [S]everal large English tests hold sway Global-wide; tests which are clear agents of the English variety of the nation where they are produced. These tests maintain their agency through the statistical epistemology of norm-referenced measurement of language proficiency, a very difficult beast to assail. (ibid., pp. 119-20)

However, some ELF researchers adopted a communicative-oriented approach and argued against assessing language proficiency in relation to an idealised native speaker norm (e.g., Canagarajah, 2006b; Jenkins, 2006b). For example, Canagarajah (2006a) argued that norms are "relative variable, heterogeneous, and changing" (p. 234). He further argued that the issue of which norm to adopt is no more relevant since proficiency means the ability to communicate with other speakers coming from diversity of English norms and speech communities (ibid.). In line with this, he encouraged language testers:

> [T]o shift […] emphases from language as a system to language as social practice, from grammar to pragmatics, from competence to performance. Of course, these constructs are not exclusive. However, bias in language teaching and testing circles is still very much on the first construct in each pair. Defining language use as performative involves placing an emphasis on the second construct in each pair and considering how language diversity is actively negotiated in acts of communication under changing contextual conditions (p. 234)

Similarly, Jenkins (2006b, p. 48) argued that language tests should be "communicatively motivated," and thus testers need to consider language variability something normal when assessing examinees rather than penalizing them for not conforming with standard norms, despite their success in communicating.

In response to such arguments for assessment based on ELF norms as regards L2 English language testing, L2 English language assessment has gradually changed its focus from the native

speaker model towards communicative-based assessment valuing the interactive strategies that test takers adopt for attaining communication goals (Taylor, 2006). This gradual change of L2 English language assessment included: (1) the emphasis on can-do statements; (2) the incorporation of communicative assessment tasks; and (3) shift of research focus (ibid.). The first change can be seen in the rating criteria for L2 English oral proficiency tests where performance is rated using linguistic forms besides other factors such as coherence, discourse management and interactive strategies (ibid.). Although the CEFR (Council of Europe, 2001) in general stresses native-speaker norms (as mentioned above), some rating criteria of this framework can be an example of this change. A current example is as follows (retrieved from Council of Europe on 5th February, 2022): "Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters" (level A2). These rating criteria include some can-do statements, suggesting some change of focus from native speaker model and form to function and communication. The second change can be noticed by the increase in the use of pair work to assess intercultural communication skills in several high-stake tests, which is a way to foster language testers' awareness of ELF (Taylor, 2006). As for the last change, it can be noticed by observing the focus of research on impact of accents on listening test assessment (Major et al., 2002) and its focus on assessing oral performance in communications among English NSs and NNSs (Kim, 2009).

Considering the arguments for an ELF position in L2 language assessment demonstrated above, this PhD study investigates whether EFL teachers' attitudes are oriented to ELF intelligibility in pronunciation assessment, or they still focus on native speaker norms in pronunciation assessment (i.e., nativeness; as mentioned in Chapter 2; see section 3.4 below for more details on ELF intelligibility and more arguments for it). If the EFL teachers in this PhD study are oriented to ELF intelligibility, this indicates they adopt the justified ELF position in pronunciation assessment, but if they, otherwise, are oriented to nativeness, this will necessitate training and education for these teachers to change their attitudes to align with the ELF position.

3.4 A Lingua Franca Approach to L2 English Intelligibility

3.4.1 Definition of the Lingua Franca Approach

The lingua franca approach focuses attention on communication instead of idealized/NS norms (Sannes, 2013). It is a flexible approach that reflects the cultural norms of all users of English (Kirkpatrick, 2006), not only those of NSs. This approach is an alternative that can be "liberating"

since it suggests that English is in the ownership of all its speakers, rather than NSs only (Sannes, 2013, p. 27).

A lingua franca approach to L2 intelligibility focuses on transnational interactions, and adopts intelligibility, rather than nativeness, as a basis for the approach (Sewell, 2017). It aims to accommodate the variations among L2 varieties and by extension the variations among L1 varieties (ibid.); it suggests that such variation should be accepted (ibid.). As stated by Kirkpatrick (2006), "Communication in lingua franca English has to be seen in terms of accommodation between codes and in a multilingual context" (p. 80). Similarly, Jenkins (2000, 2002) has attached critical importance to the accommodation among ELF users towards each other, which demonstrated great tolerance for variation. Therefore, a lingua franca approach to L2 intelligibility comes to prioritize adaptability and flexibility over the ability to reproduce a predefined system—e.g., to mimic NS norms— and, at the same time, it recognizes the continuing importance of intelligibility (Sewell, 2017).

3.4.2 Practicality of a Lingua Franca Approach to L2 English Intelligibility

The ideology of native speakerism has reinforced the idea that English belongs to the NSs of a given language (Monfared & Safarzadeh, 2014). For a long time, British and American Englishes have been predominantly considered the acceptable standards for teaching and assessing English around the global, specifically in expanding circle countries (ibid.). Previously, learning English as a second or foreign language generally followed RP or GA as a model, and NNSs are expected to try to imitate NSs in order to be easily understood or to attain assessment goals (Deterding & Gardiner, 2018). However, the development of English as a lingua franca has raised doubt over the blind acceptance of RP and GA (Graddol, 2006), and it has given priority to the acceptability of NNS varieties of English (Jenkins, 2007; Kirkpatrick, 2010a; Seidlhofer, 2004). According to Deterding and Gardiner (2018), a large number of interactions in English nowadays occur between NNSs but do not involve NSs; thus, adopting a NS English model for L2 English pedagogy is increasingly observed as irrelevant (ibid.). This is not just because a large proportion of NNSs do not have the desire to sound like native speakers, but also because some of the NS features do not contribute to intelligibility in the global context (Smith & Nelson, 1985). An early study conducted by Smith and Rafiqzad (1979) investigated the English intelligibility of English NSs and NNSs coming from 10 countries (Hong Kong, India, Japan, South Korea, Malaysia, Nepal, the Philippines, Sri Lanka, Thailand and the USA). The findings show that the NSs from the USA were the least intelligible

speakers according to the ratings of L2 English listeners coming from 11 countries (Bangladesh, China, Hong Kong, India, Indonesia, Japan, South Korea, Malaysia, Nepal, the Philippines and Thailand) (ibid.). The authors of this study concluded that NSs are not necessarily more intelligible than NNSs (ibid.).

English now serves as a global lingua franca that connects the world (Crystal, 2003) and facilitates cultural understanding across societies (Phan Le Ha, 2008). As Kachru and Nelson (2001) stated, "today, English is spoken and taught worldwide and the language is used primarily by non-native speakers to communicate with non-native speakers" (p. 13). So, it makes no sense to enforce particular communication norms or monolingual and mono-cultural norms as an ideal model in the field of ELT (Phan Le Ha, 2008).

According to Moussu and Llurda (2008), not only do NNSs outnumber NSs of English, but also NNS EFL teachers outnumber NS EFL teachers. According to Canagarajah (2004), the approximate percentage of NNS L2 English teachers is 80%. Collectively, the approximate number of English users worldwide is 2 billion users (Schneider, 2011). This suggests the existence of a great diversity of English varieties which has led to an increasing research interest in the linguistic features essential for intelligibility of all these varieties (ibid.). Accordingly, Lee (2012) stated:

> When English is learned by millions of bilingual speakers as an additional language for international communication, it is necessarily denationalized and acculturated to local specific needs. Hence it is unacceptable that NS-based norms should prevail and serve as the yardstick for measuring NNSs' phonological accuracy, lexico-grammatical correctness and discourse-pragmatic appropriacy. (p. 9)

The practicality of applying a lingua franca approach to L2 English intelligibility can further be supported by similar arguments that did not address the practicality of this approach directly. For example, Labov (1969, as cited in Cook, 1999) argued that one group should not be assessed in relation to another group's norms, whether these two groups are different based on class, race, sex or any other categories. He further argued it is not anticipated that people follow the norms of other people they do not belong to (ibid.). He further argued that even though nearly all teachers and researchers would agree that the differences between groups of people do not make one group inferior to another, many of these teachers and researchers, and of people in general, have frequently considered L2 learners differently (ibid.); they have considered L2 learners a group that

can be measured against the norms of the NS group (ibid.). Labov further mentioned that this invalid comparison is similar to the one made between women and men: it is wrongly claimed that women should speak like men to be successful in business. It is also similar to the one made between children from different ethnic and/or social backgrounds: it is wrongly claimed that black children should be educated to speak like white children and the working-class children should be educated the language of those of the middle class (ibid.). In addition to this, Housen and Kuiken (2009) questioned the relevance of L2 performance assessment that is based on NS norms, and they instead recommended the use of non-standard/NNS norms which are more practical and being progressively used in many contexts. In the field of second language testing, the research has developed to include debates about practices and theories that reflect the re-orientation of the English language suggested by research within sociolinguistics (Hsu, 2012). This research has addressed the "function, status, linguistic maturity and legitimacy of the multiple varieties of English indicating that English language can and should no longer be viewed as a homogenous entity" (ibid., p. 1). For example, the international language test provider "Cambridge ESOL" has reconstructed its practices and policies following the notion that L2 communication in the global context could still occur regardless of the multiple varieties of English (Taylor, 2002, as cited in Hsu, 2012). Furthermore, according to Matsuda and Friedrich (2011, p. 332), traditional approaches to EFL instruction, which promote nativelikeness, could not "adequately prepare" L2 learners of English to effectively communicate with speakers "from other English-speaking contexts." Thus, a lingua franca approach to L2 English intelligibility, rather than intelligibility based on NS norms, is the ideal way to address the issue of diversities of English used worldwide by both English NNSs and NSs.

3.4.3 A Lingua Franca Approach to a Model for L2 English Pedagogy

The main interest of ELF research is identifying salient common features for the use of English, regardless of speakers' L1s and their L2 proficiency levels (Seidlhofer, 2004). Seidlhofer (2006) argued that in order to provide L2 learners of English with an alternative model, we need to analyse how English is spoken as a lingua franca. She further argued that we need to analyse ELF corpora to identify the rules ELF speakers follow in their communications to understand each other (ibid.).

After rejecting NS-based models, Alptekin (2002) called for an urgent look for a new model based on a lingua franca approach to teaching English that takes into account the international

position of the English language. Accordingly, Alptekin (ibid.) suggested that such a model should consider the following criteria:

> "1) Successful bilinguals with intercultural insights and knowledge should serve as pedagogic models in [ELF] rather than the monolingual native speaker;
>
> 2) Intercultural communicative competence should be developed among [ELF] learners by equipping them with linguistic and cultural behavior which will enable them with an awareness of difference, and with strategies for coping with such difference […];
>
> 3) The [ELF] pedagogy should be one of global appropriacy and local appropriation, in that it should prepare learners 'to be both global and local speakers of English and to feel at home in both international and national cultures' […];
>
> 4) Instructional materials and activities should involve local and international contexts that are familiar and relevant to language learners' lives;
>
> 5) Instructional materials and activities should have suitable discourse samples pertaining to native and nonnative speaker interactions, as well as nonnative and nonnative speaker interactions. Discourse displaying exclusive native speaker use should be kept to a minimum, as it is chiefly irrelevant for many learners in terms of potential use in authentic settings […]." (p. 63)

3.4.4 A Lingua Franca Approach to L2 English Pronunciation Intelligibility

ELF research has aimed attention at three aspects of language: pronunciation, pragmatics and lexico-grammar (Seidlhofer, 2004). According to Hsu (2012), of these three aspects, a great stress is laid on pronunciation. Monfared and Safarzadeh (2014) suggested that, the concept of pronunciation is one of the most challenging issues in the area of ELF, and thus this justifies the great focus on pronunciation. Pronunciation has been greatly discussed by many scholars over the last decade or so (e.g., Jenkins, 1998, 2000, 2002, 2004, 2009a, 2009b; Derwing & Munro, 1997, 2005), with Jenkins being the lead proponent for the lingua franca approach. In her discussion of pronunciation, Jenkins (2002) argued that the intuitions that are taken into consideration as regards pronunciation are those of NSs and little attention is given to the intelligibility for NNSs, despite the fact that NNSs significantly outnumber NSs. Jenkins (2000, p. 207) also pointed out, "a

native-like accent is not necessary for intelligibility in ELF interaction". Also, Jenkins (1998, 2000, 2002, 2004) has placed great stress on ELF pronunciation and its pedagogical issues. For example, Jenkins (1998) stated that:

> While approximation to the native model is probably essential for intelligibility in non-bilingual [ELF] contexts as regards core sounds, nuclear stress, and relevant articulatory setting, local non-native norms are likely to be both acceptable and intelligible in many other phonological areas. (p. 124)

Further, according to Jenkins (2009c), all English speakers, including NSs, are members of ELF community. She pointed out that when the members of Inner Circle community participate in ELF communication, they do not set the linguistic agenda (ibid.). From an ELF perspective, all speakers of English, including NSs, need to make fine-tuning to their local English varieties in order to assist their interlocutors during communication (ibid.). According to Khatib and Monfared (2017, p. 215), the great emphasis given to pronunciation in ELF research could be attributed to two main causes: first, "constructing intelligibility which refers to creating a comprehensible discourse among participants within a communicative framework"; and second, "taking into consideration the principal role of pronunciation in [ELF] and the way that it shapes learners' awareness towards their sociocultural identity".

A main text in the development of a lingua franca approach to L2 English intelligibility pronunciation is Jenkins' (2000) *The Phonology of English as an International Language*. Hsu (2012) asserts that Jenkins (2000) is one of the first to study intelligibility in oral interactions between NNSs of English. According to Walker (2010), Jenkins' (2000) text can be considered as a model for teaching and assessing L2 pronunciation. In the process of writing her book, Jenkins (ibid.) argued that English NNSs who have different L1s probably rely on bottom-up strategies (i.e., reliance on acoustic information) to interpret other NNSs' accented speech instead of top-down strategies (i.e., dependence on contextual information). Thus, Jenkins (ibid.) approached the term "intelligibility" in the spirit of researchers such as Bansal (1990), where, accordingly, successful oral communication is based on correct use of phonology, not only in international contexts, but in intra-national ones as well. For example, Bansal (1990) identified a number of phonetic features which are inclined to affect the intelligibility of spoken English in India such as the following: lack of clear articulation, vowel or consonant substitution, and stress on the wrong syllable. As a result, Jenkins (2000)

argued that intelligibility entails the production and recognition of words and utterances and, especially, the ability to produce and receive phonological form.

Jenkins (2009a) reported that she had carried out a large-scale empirical research project to identify which features of RP or GA are crucial for intelligibility in ELF communication and which are not required for—or even detrimental to—intelligibility. According to Jenkins (ibid.), she collected the data from NNSs of English with a large number of first languages interacting orally with each other in a wide range of contexts, both educational and social, and then she analysed the data to identify which intelligibility features that could be associated with pronunciation. The items that emerged as necessary for intelligibility she labelled the Lingua Franca Core (LFC; ibid.). Jenkins (2009b) stated that the LFC is meant for oral communication on a global level rather than with NSs only (ibid.). It is supposed to promote better intelligibility among ELF interlocutors than many NS varieties (ibid.). Finally, the LFC legitimises NNS accents (Jenkins, 2005; Jenkins, 2007), and it maintains the social identities of the speakers of NNS accents (Walker, 2010; see sections 3.5.2 and 3.5.3 below for arguments for and against the LFC).

However, although the LFC is the most recognisable paradigm in current research in terms of identifying the phonological features essential for pronunciation intelligibility in international oral communication, earlier consideration of such a paradigm was proposed by Gimson in 1978 (Cruttenden, 2014). Gimson (1978) introduced a rudimentary international pronunciation model for NNSs which allowed sound modification that has only limited impact on general intelligibility. According to this model, most modifications to vowels are allowed while only slight modifications to consonants are allowed, and thus, pronunciation consonants in general should be nativelike (ibid.). This is similar to Jenkins's recommendations for segmental aspects of the LFC. Another model was suggested by Jenner (1989, as cited in Dimova, 2018), where he introduced the notion of a pronunciation core, which was later expanded by Jenkins (2000, 2000) to become the LFC. Following Jenkins' LFC, other scholars have explored the pronunciation features that can cause oral communication breakdowns in the context of ELF if they are not produced accurately. For example, some of these scholars found that some features of the LFC can cause oral communication breakdowns while some non-core features, which are not a part of the LFC (see the following section), could cause intelligibility features (e.g., Pitzl, 2005, as cited in Kim & Billington, 2018; Osimk 2009, as cited in Kim & Billington, 2018). Overall, the LFC is still the most recognisable paradigm in terms of the phonological features essential for L2 English pronunciation intelligibility.

It has received much investigation and discussion; most of this investigation and discussion has given it support (see section 3.5 below).

3.5 The Lingua Franca Core

In this section, details of the LFC that were extracted from Jenkins (2000) will be presented. Firstly, the LFC includes both segmentals and suprasegmentals. As regards the segmentals, there are 24 consonant sounds in RP and GA in common vital for L2 pronunciation intelligibility. The LFC requires, as a rule, EFL learners to produce these consonant sounds in a way, at least, similar to that of RP and GA, rather than to imitate these varieties. However, there are some other rules regarding consonant sounds according to the LFC:

- The dental fricative pair /θ/ and /ð/

The dental fricative pair /θ/ and /ð/ are entitled to substitution according to the LFC; substitution of these phonemes do not lead to pronunciation unintelligibility.

- The dark /l/

Another omission from the LFC is related to a phonetic feature rather than phonemic: the substitution of dark /l/ with regular substitutions such as clear /l/ or a vowel similar to /ʊ/ is acceptable according to the LFC.

- /r/

Where RP and GA differ, the GA rhotic variant, which is the retroflex approximant [ɻ], is opted for in the LFC, rather than the RP post-alveolar approximant [ɹ].

- The intervocalic /t/

The LFC follows RP in terms of the consonant /t/ once it occurs intervocalically, in contrast to the GA use where it becomes the voiced flap [ɾ].

- The fortis plosives /p/, /t/ and /k/

The LFC includes the phonetic features: aspiration [ʰ] following the fortis plosives /p/, /t/ and /k/ once they take place at the beginning of a stressed syllable; and the shortening effect of the fortis plosives on a vowel sound preceding them in contrast to the lenis consonants /b/, /d/ and /g/ which maintain the length of a preceding vowel sound.

- Insertion of sounds

Insertion of sounds into a consonant cluster can sometimes be problematic for intelligibility according to the LFC, especially in the case of epenthesis in a stressed syllable (e.g., the pronunciation of the word "stroke" as ['sɪtɹəʊk]), or where the insertion creates a homonym (e.g., the pronunciation of the word "hard" as ['hɑːdə], which sounds like "harder").

- Consonant cluster simplification

Consonant cluster simplification is acceptable according to the LFC in the case of consonant elision, but not in the case of consonant deletion. Consonant elision is simplification of a consonant cluster governed by rules of English (e.g., elision of the sound /d/ in the word "friendship"). In contrast, consonant deletion is governed by the constraints of syllable structures in learners' L1s (e.g., elision of the sound /r/ in the word "product"). The situation for elision is the same for both RP and GA, except for the intervocalic cluster "-nt-" when it occurs before an unstressed syllable (e.g., the cluster "-nt-" in the word "twenty"); following the RP model, the LFC does not allow for elision of /t/.

There are two main considerations as regards vowel sounds in the LFC: quality and quantity. Vowel quality is concerned with tongue and lip position, while vowel quantity is concerned with length. Following are the rules for vowels according to the LFC:

- Vowel quantity

Vowel quantity contrasts must be preserved according to the LFC, e.g., "pill" versus "peel".

- Vowel quality

The actual quality of vowels is less important than vowel quantity, providing an L2 speaker is consistent, e.g., the speaker must not switch between different pronunciations of the vowel in the word "hat" so sometimes it sounds like RP /hæt/ and sometimes it sounds like New Zealand /het/.

- The mid-central unrounded vowel sound /ɜ:/

The quality of the mid-central unrounded vowel sound /ɜ:/ must be preserved since Jenkins's (2009a) data (of her large-scale empirical research project; see section 3.4.4 above) indicated that the change of the quality of this vowel induced unintelligibility.

- Diphthongs

The arguments regarding vowel quantity and quality also hold true for diphthongs: their length contrasts must be maintained, and their qualities must be used consistently regardless of which qualities an L2 speaker uses.

As for suprasegmentals, the LFC has the following rules:

- Word stress

Word stress is not essential for the intelligibility of individual words according to the LFC but, since it has implications for nuclear stress and sound identification, the LFC recommends teaching it to L2 English learners in form of general guidelines, while noting the many exceptions of word stress rules.

- Intonation

The shape of intonational contour (rising, falling, etc.) is not essential for intelligibility according to the LFC.

- Nuclear stress

Nuclear stress is the most important key to the speaker's intended meaning in an utterance, as it is the salient part that indicates where the listener should pay attention. As such, it is considered

essential for intelligibility in the LFC. According to the LFC, English speakers should include one nuclear syllable (although complex word groups include more than one nucleus) in each meaningful unit (or chunk) of their utterances, whether the nuclear stress is unmarked (on the last content word in a word group[1]) or contrastive (somewhere else; see Table 3.1 below for intelligible, according to the LFC, and unintelligible examples of the two different types of the nuclear stress).

Table 3.1: Intelligible and unintelligible examples of the different types of the nuclear stress

Nuclear stress type	Example	Intelligibility
Unmarked	A. What do you want to buy? B. I want to buy a lemon.	Intelligible
-	A. What is about to collapse? B. The bridge is about to collapse	Unintelligible
Contrastive	A. You took my place. B. I didn't take your place.	Intelligible
-	A. You're busy. B. I'm not busy.	Unintelligible

3.5.1 Empirical Research into the Lingua Franca Core

As far as the author of this study is aware, there are remarkably few studies replicating Jenkins's original study on the LFC (Jenkins et al., 2011) or investigating the LFC (Jenkins et al., 2011), and none which consider its role in attitudes to pronunciation assessment. According to Jenkins et al. (2011), this limitation of research may be because attention has shifted from:

> [A]n orientation to features and the ultimate aim of some kind of codification (an aim which, nevertheless, has not been dismissed out of hand), to an interest in the processes underlying and determining the choice of features used in any given ELF interaction. (p. 287).

[1] A word group are words which together form a meaningful unit, and it is separated from the preceding and/or following word group(s) by a pause at the boundaries or, less commonly, by a change in an overall pitch level or rhythm (Jenkins, 2000).

The majority of the findings in the research replicating or investigating Jenkins's original study mostly support the LFC (e.g. Da-Silva, 1998; Deterding, 2011, 2013; Matsumoto, 2011; Osimk, 2009; Zoghbor, 2011). For instance, Osimk (2009) concentrated on the following features: aspiration, realisations of /θ/ (e.g., in the word "bath") and /ð/ (e.g., in the word "bathe") and /r/. The focus was given to /θ/ and /ð/ since, despite given much attention in classrooms, the LFC states that accurate realisation of these two sounds is inessential for NNS-NNS interactions. The findings of Osimk's (2009) study as regards the two dental sounds and aspiration support the LFC, where they indicate the significance of aspiration for mutual intelligibility and showing that the two sounds /θ/ and /ð/ are unimportant. However, the findings as regards rhotic /r/ seem less conclusive (ibid.). Also, Matsumoto (2011) analysed interactional data of six ELF speakers and found that the deletion or mispronunciation of consonants except for the dark /l/ and the dentals /θ/ and /ð/ led to intelligibility problems. This is consistent with the LFC proposals that consonants except for the dark /l/ and the dentals /θ/ and /ð/ must be preserved.

Da Silva (1998) also investigated the LFC features. He conducted an experiment where he had three of his Brazilian students communicate with twelve students from nine different countries speaking nine different languages (ibid.). The main findings of his study show support for the LFC. For example, he found that the sounds /θ/ and /ð/ were unimportant for intelligibility (ibid.), supporting the LFC in this regard. However, Da Silva also found that some mistakes in the features included in the LFC did not seem to impair intelligibility (ibid.). For example, some of the students in the study shortened some long vowels in the words "read" and "speak", but that did not seem to impair intelligibility (ibid.).

Another example is Deterding's (2013) corpus-based study. The overall conclusion of this study is that the great majority of the tokens of misunderstandings in spoken ELF as far as phonology is concerned came from the following: (a) omitting consonants from initial or medial clusters, (b) replacement of or dropping consonants except for the dark /l/ and the dentals /θ/ and /ð/ (ibid.), (c) mispronunciation of the mid-central unrounded vowel sound /ɜː/, (d) wrong production of vowel quantity, and (e) misplacement of nuclear stress. These findings are consistent with the LFC proposals. However, Deterding (ibid.) found that a few tokens of misunderstandings in spoken ELF (as far as phonology is concerned) came from the following: (a) replacement of the dark /l/, (b) mispronunciation of the dentals /θ/ and /ð/, (c) omitting consonants in final clusters, and (d) misplacement of word stress.

Deterding's (2011) corpus-based study also shows that most of the misunderstandings of spoken ELF were caused by consonants. In the same study, Deterding (ibid.) found that the mid-central unrounded vowel sound /ɜː/ was also important for intelligibility in spoken ELF, which is also consistent with the LFC proposals. However, Deterding (ibid.) found that /ð/ might have affected intelligibility in spoken ELF, which is inconsistent with the LFC proposal that /ð/ is not essential for intelligibility (see section 3.5 above). Also, Zoghbor (2011) found similar results when she compared the performance of L2 English learners who followed two different pronunciation syllabuses, where one syllabus was based on the LFC and the other on more traditional approaches, but her findings are not statistically significant.

Pickering (2009) found some results supporting the LFC and others contradicting it. Pickering (ibid.) analysed a 17-hour data set from 25 NNSs of different proficiency levels and of different L1 backgrounds. These participants completed the following interactive tasks: direction giving, information gap, dialogue reading and informal conversation. The findings of Pickering's (ibid.) study demonstrate the significance of nuclear stress; Pickering (ibid.) found that nuclear stress was important for intelligibility in ELF interactions. This supports the importance given by the LFC to nuclear stress. However, Pickering (ibid.) found that intonation contributed to intelligibility, contradicting the LFC. She found that rising tones were used when an interlocutor approached an unknown lexical item, and falling tones were used to confirm understanding and the negotiation has been achieved successfully. She also found that pitch height was used for signalling a feature of understanding and suggesting for a repair in case of a high pitch and signalling agreement in case of a mid-pitch. Examples of the high pitch and mid pitch from Pickering (2009, p. 247) are shown below, where → refers to a mid pitch and ↑ to a high pitch (see Appendix C for full pitch transcription conventions).

(1). Taiwanese L1 speaker (T) and Korean L1 speaker (K)

T: //→ ok LEAVES leaves of PLANT //

K: //→ YEAH//

T: //→ how many LEAVES //

K: //→ ONE just one//

T: // ↑ just ONE // ↑ LEAFS//

K: //→ uh//

T: //↑ plant LEAFS // ↑ just ONE//

However, Field (2005) found results contradictory to the LFC proposals regarding the importance of word stress (as opposed to nuclear stress in sentences/chunks). In his study, Field (ibid.) included 24 disyllabic English words. These words were recorded with a word stress in two places: the standard place and shifted from its standard place to a wrong one. The recorded words were then presented as single words rather than embedded in sentences. These words were played to L2 English learners at two language schools, and they were asked to transcribe them. The results of Field's (ibid.) study show that the shift of word stress from its standard place in a word reduced intelligibility considerably among the L2 English learners. However, even though stress is not a core part of the LFC, Jenkins (2000) still recommended teaching it as it has implications for nuclear stress (see section 3.5 above).

Although some of the studies mentioned above give support to the LFC, they have just investigated a small part of the LFC (e.g., Da Silva, 1998; Matsumoto, 2011; Osimk, 2009; Pickering, 2009). The LFC has addressed many aspects of phonology and phonetics, but these studies have just investigated few features of the LFC. Some other studies have investigated all the features of the LFC, but their findings still do not support it completely (e.g., Deterding, 2011, 2013; Zoghbor, 2011). Given all this, the LFC cannot be proved completely valid at this point until more research is conducted to investigate it. The current PhD study has not investigated the validity of the LFC since it was not a main purpose of the study, and it was beyond its scope, but it is discussed briefly in the discussion chapter (Chapter 7; see section 3.5.3 below for a justification for adopting the LFC as a framework in this study).

3.5.2 Arguments for the Lingua Franca Core

In addition to the empirical studies that support the LFC (see section 3.5.1 above), there are arguments in the literature supporting the LFC, but there are only a few (as far as the author of this study aware). For example, Walker (2001) discussed the responsibility of EFL teachers to change the priorities of learning English from the attempt to achieve nativelike competence to trying to achieve an acceptable level of intelligibility. In this regard, he argued for the LFC, and emphasised that this is the only approach to pronunciation intelligibility based on empirical evidence (ibid.). Walker (ibid.) further argued that an effective solution for identifying priorities for intelligible pronunciation is to conduct a contrastive analysis between the phonological systems of L1 and L2. Accordingly, the LFC proves advantageous as it was informed by contrastive analysis (ibid.). This in return has two advantages to L2 English teaching and learning: "a) the total workload required of teacher and learner is now greatly reduced; b) the new goals are more achievable both in terms of teaching and learning" (Walker, ibid., p. 2). Finally, Walker (ibid.) argued that adopting the LFC as a framework for intelligible pronunciation in L2 English would lighten the psychological burden on L2 English learners by focusing on achievable goals in learning L2 English rather than impossible goals for attaining nativelike competence.

Timmis (2018, p. 501) summarized the pedagogic arguments in favour of the LFC, as follows: "the LFC presents a lighter and more focused pronunciation workload than the traditional procession towards native speaker norms by focusing on what is teachable and learnable and of proven value in NNES-NNES intelligibility". He also argued that learners who are provided with such a paradigm or model are "faced with a manageable and practical" paradigm and thus they "will gain in motivation and in self-esteem and probably feel empowered to express their own identity through their accent" (ibid., p. 501).

Moreover, the LFC can help in maintaining "uniformity" of ELF where NNSs who come with diverse linguistic backgrounds could have one common and more attainable aim for pronunciation (Khatib & Monfared, 2017, p. 218). It can further assist in developing L2 English learners' "awareness" of present globalization of English in order to be more "realistic" about their learning goals and not to be "just blind followers of a particular model" (ibid., 214).

3.5.3 Arguments against the Lingua Franca Core

Although the LFC could serve as a syllabus that advises L2 English instructors and learners, the validity of this core is not without features. There are some arguments in the literature against the LFC. For example, one argument suggests that the LFC might cause diversification in language use, and thus it may eventually lead to unintelligible varieties (Dziubalska-Kołaczyk, 2005; Yamaguchi, 2002). However, Jenkins (2000) argued that this is unlikely to occur since, according to the language universals theory (Anderson, 1987; Jakobson, 1941), there is a universality of solutions and substitution of sounds that are used by interlocutors when L2 English features do not exist in L1 (see, for example, section 3.5 above). This is also one of the reasons behind Jenkins's (2000; 2002) suggestion that learners of ELF pronunciation should be exposed to speakers of English with different L1s (along with NSs) in order that they recognize the alternatives used by those speakers for some phonemes.

Other arguments concern the grounds on which Jenkins selected the features of the LFC. For example, Derwing (2008, p. 352) mentioned that "the available evidence is very limited, based on a small sample of communication breakdowns." Similarly, Dauer (2005) commented that:

> It is important to note that Jenkins' data are based on a small number of well-educated, motivated [NNSs] whose proficiency in written English, which they learned first, is at the intermediate to advanced level […] It is not clear whether these results can be generalized to a larger population of less literate learners or to more formal extended discourse. (p. 549)

Similarly, Isaacs (2018) stated that the LFC was developed based on a limited data drawn from the interactions between few international students in England. Thus, it is problematic to generalise the LFC to all international contexts where English is used for communication as this "likely overstates the case" (ibid., p. 282). However, as mentioned above in section 3.4.4 above, Jenkins (2009a) reported that she had carried out a large-scale empirical research project to develop the LFC. She reported collecting the data from NNSs of English with a large number of first languages interacting orally in a wide range of contexts, both educational and social (ibid.).

Isaacs (2018) further argued that there is a lack of a comprehensive reporting of the data collection for the LFC, and there is a lack of description of the tasks used. Similarly, enough detail on the data collection methodology and the tasks and the error types were not found by the researcher

of this PhD study in any text by Jenkins. Isaacs (ibid.) further stated that, in order to adopt the LFC and use it as a standard for pronunciation teaching and assessment, there is still a need for further empirical evidence and "validation work" for the core to replace the NS standard (ibid., p. 282).

There are also some specific reservations expressed by Dauer (2005, pp. 547-548) as regards the LFC. These reservations are as follows:

- The recommended substitution of /θ/ and /ð/ with /f/ and /v/ is not useful for the many learners who have a feature with /v/. Thus, /t/ and /d/ might be better substitutions.
- The omission of word stress is inappropriate since some aspects of the LFC are closely related to word stress. As Dauer (2005) commented:

  > It is hard to understand how to teach aspiration, vowel length, or nuclear stress (all of which are part of the LFC and are associated with word stress) without students having been taught which syllable to stress in a word. (p. 548)

However, these reservations seem to be based on misunderstanding of the LFC. The LFC did not recommend substitution of /θ/ and /ð/ with /f/ and /v/; instead, according to the LFC, /θ/ and /ð/ are commonly substituted by a restricted set of alternatives — i.e., /t/ and /d/, /s/ and /z/, or less commonly /f/ and /v/ (see section 3.5 above for more details on this argument and the following two arguments in this paragraph). Furthermore, the LFC did not recommend the omission of word stress; instead, according to the LFC, word stress is not essential for the intelligibility of individual words but, since it has implications for nuclear stress and sound identification, it is recommended to teach it to learners in the form of general guidelines.

Further arguments against the LFC are presented by Riney et al. (2000); they argued that it is still not clear how to design a phonological syllabus that is based on prioritized phonological features for a particular L2 English context where both global and L1-related pronunciation features affect speech intelligibility. Similarly, Rugesæter (2012, as cited in Sannes, 2013) cast doubt on a global model such as the LFC and confirmed that pronunciation teaching should be conducted while taking each L1 background into account separately. He further argued that because of differences in L1s competences, L2 speakers with a distinct L1 backgrounds encounter difficulties in different L2 aspects, and thus any global model for intelligible pronunciation could not tackle this feature (ibid.). For example, the dearth of voiced fricatives in the Norwegian language could create a difficulty with discrimination between sounds like /s/ and /z/; however, those speakers may not face a problem

with sounds like /l/ and /r/ which could be problematic for Chinese and Japanese L2 learners of English as those two sounds do not exist in the languages of both groups of learners (ibid.). However, to develop the LFC, Jenkins (2009a) reported that she depended on interactions between NNSs with a large number of L1s, thus it seems that the LFC would help improve interactions between NNSs with such many L1s at least. Also, if a global model for each NNSs' interaction is impossible, the LFC can at least be a general description of the phonological features essential for the intelligibility of specific ELF interactions, but it can be developed later to include broader contexts.

A final argument against the LFC is offered by Prodromou (2006), where he stated that advanced L2 English learners will always think that there is a more prestigious and expert level of proficiency in L2 English among English NSs that is superior to an ELF model regardless how effective that ELF model is (ibid.). However, in the view of the author of this PhD study, the purpose of the LFC is not to make L2 learners think that this model is a standard and they should not be aspired to achieve nativelike competence. Instead, the LFC is a working model that aims at providing features essential for intelligible communication between the users of English, and the argument that some L2 English learners will think of a more prestigious and expert level seems irrelevant to the purpose of the LFC. Anyway, L2 learners should be educated that nativeness is almost unachievable by L2 learners (Abrahamsson & Hyltenstam, 2009; Jenkins, 2000; Loewen, 2015; Moyer, 1999; Munro & Derwing, 2011; Singleton, 2005), and there is the problem of determining who is a NS (Cook, 1999), as there is variation within native varieties (Lindemann, 2017). As a result, L2 learners may want to focus on being intelligible rather than aiming for nativelike competence.

Despite the issues raised above, the LFC can be considered a working model for teaching pronunciation (Khatib & Monfared, 2017), at least for beginning stages. It can help both L2 English teachers and learners with identifying the features that are important for intelligible pronunciation (ibid.). Also, according to Isaacs (2018, p. 281), the LFC is a working model but "it needs to be used with caution and in conjunction with additional research evidence on what counts the most for intelligibility." Similarly, Walker (2010, p. 44) recommended the use of the LFC, but he further argued that there are still some areas of this core that require a degree of "fine-tuning."

For the moment, the LFC can provide a solution to choice of model, as it normalises the variations in English pronunciation in reference to accents such as RP and GA and in other English

NNS accents to which learners might be exposed. However, the LFC might be used as a general guidance only with the ability to introduce minor modification to it as mentioned above.

3.5.4 The Lingua Franca Core as a Model of Pronunciation

A related discussion to the LFC is whether this core is considered a model of pronunciation. The LFC has been considered by some researchers as a model of pronunciation (e.g., Dauer, 2005; Kirkpatrick, 2006). However, in her discussion of the LFC, Jenkins (2000) characterized the LFC as a set of instructional priorities intended to facilitate intelligibility between NSs and NNSs of English, but she argued against the categorization of the LFC as a model since there is no accent of English which includes all and only the LFC features. In support of this, an earlier consideration of such a paradigm, which was proposed by Gimson in 1978 (Cruttenden, 2014), proposed a set of pronunciation priorities for learners, but it was not suggested as a model for imitation or approximation.

In line with this, Jolanta (2018) argued that the concept of a model is usually confused with the aim of pronunciation teaching and learning. This confusion is caused by the belief that the aim of pronunciation teaching and learning is to master all the phonetic features of a model accent (ibid.). However, assuming that not all model features need to be acquired, this will mean that with one model different pronunciation teaching and learning goals can be pursued (ibid.). Thus, a particular pronunciation model does not determine the pronunciation learning goals an L2 learner tries to achieve (ibid.). Rogerson-Revell (2011) similarly argued that:

> There is a difference between the reference points or model we use for guidance in pronunciation teaching and learning and the target we set as an achievable goal or aim. The 'goal' is the level which a learner's pronunciation aims to reach in order to facilitate effective communication. (p. 8)

It is the view of the author of this PhD study that the LFC should not be considered a model, but it should be used as guidance for developing a syllabus for pronunciation teaching and assessment (as mentioned above; see section 3.5.3), such as the case of Zoghbor's (2009) syllabus for pronunciation (see below).

3.5.5 An English Pronunciation Syllabus Based on the Lingua Franca Core

Zoghbor (2018) tried to create a syllabus for English pronunciation based on the LFC for Arab learners of English. According to Zoghbor (ibid.), traditionally, L2 English pronunciation has often been compared with the phonology of the most common native varieties of English, RP and GA. However, Walker (2010) explained how the LFC can be contextualized for L2 English learners by comparing their L1 with the LFC, since comparing L2 English learners' pronunciation of English with those of RP and GA is not anymore valid comparison. As a result, Zoghbor (2009) developed an L2 English pronunciation syllabus based on the LFC for Arab learners of L2 English by making contrastive analysis (CA) between the LFC and Modern Standard Arabic (MSA). The Arabic language is used in nearly 23 countries, and in every country, there are (at least) two major varieties of Arabic: MSA and Non-Standard Arabic (NSA) (Mahmoud, 2000). MSA is a simplified form of Classical Standard Arabic (CSA), the language of the Quran (the main religious text in Islam, believed by Muslims to be a revelation from God), and it is the language for all official communications in the Arab-speaking countries, specifically in media (Mahmoud, 2000; Swan & Smith, 2001; Yorkey, 1974). MSA is the only formal written version of Arabic; in contrast, the other varieties of Arabic used in the Arab world are mainly spoken, but they may be used in written forms specifically in informal media such as social media (Zoghbor, 2018). MSA is also taught in schools as a separate subject in the Arab world (ibid.).

The pronunciation syllabus developed by Zoghbor (2009) can be found below in Table 4.1. In this table the left-hand column includes the core features of the LFC. The middle column includes the phonemes of the MSA which are obtained from Newman (2002), Swan and Smith (2001), Watson (2002), Avery and Ehrlich (1992), Kharma and Hajjaj (1997) and Kenworthy (1987). The right-hand column is the result of the CA that Zoghbor (2009) made between the core features of the LFC (the left-hand column) and the phonemes of the MSA (the middle column), and thus it is the proposed pronunciation syllabus for Arab learners.

Table 4.1: The features of the LFC, the phonemes of the MSA and the features of the pronunciation syllabus for Arab learners obtained from the CA between the LFC and the MSA (adopted from Zoghbor, 2009, p. 26).

Features of the LFC	Phonemes of the MSA	Features of the pronunciation syllabus for Arab learners
The consonant inventory: All sounds except /θ/, /ð/ and dark (or velarized) /l/ (as in 'little')	**The following do not exist in MSA:** /p/, /v/, /dʒ/, /g/ and /ŋ/. [p] exists as an allophone to /b/ before voiceless consonant phonemes (e.g. /katabt/) [ŋ] is an allophone to/n/ /l/ is alveolar. Velarized [ɬ] is an allophone of /l/	/p/, /v/, /dʒ/, /g/ and /ŋ/.
Rhotic/r/ rather than the other varieties of /r/.	Trilled /r/ (produced by vibrations between the articulator and the place of articulation)	Rhotic /r/
Intervocalic [t]	Dental /t/	Intervocalic [t]
Phonetic requirements: Aspiration after /p/, /t/, and /k/.	Aspirated [t] and [k] are allophones of /t/ and /k/.	Aspiration after /p/, /t/, and /k/.
Shortening of vowel sounds before fortis and maintenance of length before lenis consonant phonemes. i.e. /iː/ is shorter in 'seat' than in 'seed'.		Appropriate vowel length before fortis and lenis.
Avoiding contracted and short forms.	Arabs tend to avoid contracted and short forms and elisions and read with a rather heavy staccato rhythm.	Avoiding contracted and short forms.
Consonant cluster: Word initially, word medially	Arabic has no clusters of more than two consonant phonemes. Cluster does not exist in word-initial position.	Word initially and medially.
Vowel sounds: Long-short contrast	3 short vowel phonemes: /I/, /ʊ/, /æ/ 3 long vowel phonemes: /uː/, /iː/, /ɑː/ 2 diphthongs:/eɪ/, /aʊ/	Long-short contrast
Vowel quality: L2 (consistent) regional qualities. /ɜː/ to be preserved		Learners' regional quality is accepted. /ɜː/
Nuclear stress: Appropriate use of contrastive stress to signal meaning.		Nuclear (tonic) stress.

In the view of the author of this study as a native speaker of Arabic, this syllabus may not be a working syllabus since Arabs come with a wide range of varieties within Arabic. MSA is restricted to the official media and in some educational contexts only, but it is not spoken in daily communications. Ideally, there needs to be a pronunciation syllabus based on the LFC for every speaking community within the Arab world. However, like the LFC, this syllabus can be considered a working model for the moment, but it is still needed improvement in the future.

3.6 English as a Lingua Franca as a New Variety of English

Consequent to the discourse surrounding the LFC, there comes the idea of considering ELF a new variety. According to Timmis (2018, p. 500), "[p]artly as a result of the LFC proposal, confusion has arisen as to whether ELF (including grammatical, lexical and pragmatic norms) is to be regarded as a codified new variety of English or simply a mode of communication." This confusion was highlighted by Sowden (2012, pp. 90-91), where he cited Seidlhofer's (2004) remark that "ELF is a natural language and can thus be expected to undergo the same processes that affect other natural languages, especially in contact situations". In response to this, Sowden (ibid.) commented:

> This would suggest that ELF can be viewed as an identifiable, discrete entity, yet this idea is called into question by the pronouncements of other writers in the field. Jenkins (2007[, p.] 41) is clear that it does not refer to a monolithic construct which will merely replace traditional native-speaker norms with new but equally inflexible standards. Instead, it would consist of 'a variety of local versions of English', each influenced by the local native language. (pp, 90-91)

In the view of the author of this PhD study, ELF is only a mode of communication, rather than a new variety of English, since it is a description of the type of English language that speakers with different varieties use when they communicate with each other.

3.7 Summary of Chapter

This chapter has presented a review of the literature on English as a lingua franca, including the literature on the LFC. Firstly, it has addressed the development of English as lingua franca and aspects of English as a lingua franca related to the topic of this thesis. Next, the chapter has addressed the reason behind considering a lingua franca approach to L2 English intelligibility

and a major development of this approach–the LFC. Then, the empirical research and the arguments as regards this core have been presented. Next, the chapter has introduced phonological features suggested by some research as essential for L2 English intelligibility, but this research has used English NS raters as judges of L2 English intelligibility. Finally, the chapter has concluded with a related discussion about ELF.

Given the review of the literature in this chapter and the arguments introduced regarding ELF and a lingua franca approach to L2 pronunciation intelligibility, L2 English pronunciation intelligibility in this PhD study is approached from the viewpoint of ELF, and it is measured using the LFC as a premise and a framework (see section 3.5.3 above for a justification for using the LFC in this study).

CHAPTER 4: LITERATURE REVIEW 3: EFL TEACHERS' LANGUAGE ATTITUDES

4.1 Introduction

Given that EFL teachers' language attitudes are the subject of investigation in this PhD study, this chapter addresses EFL teachers' language attitudes. It begins with presenting an overview on EFL teachers' language attitudes. In this overview, EFL teachers' language attitudes are firstly defined. The overview also discusses the factors affecting EFL teachers' language attitudes since these factors are investigated in this PhD study. Next, the chapter discusses the popular methodologies used in the literature to research EFL teachers' language attitudes. This is to familiarise the reader with the methodologies used in this PhD study. In order to contextualize this PhD study and demonstrate its contribution to the literature, the chapter concludes with a review for the research looking at EFL teachers' attitudes to intelligibility and nativeness in relation to pronunciation.

4.2 EFL Teachers' Language Attitudes

4.2.1 Overview on EFL teachers' Language Attitudes

In the literature, there is no definition of EFL teachers' language attitudes in particular, as far as the author of this study knows; however, language attitudes in general have been defined by several scholars. For example, Ryan et al. (1982) defined language attitudes as "any affective, cognitive or behavioural index of evaluative reactions toward different language varieties or their speakers" (p. 7). In line with this definition, Baker (1992, pp. 12-13) suggested a multi-componential model of language attitudes that consists of "cognitive, affective and readiness for action" components. In other words, Baker's model involves "what people think about a language, how a language makes them feel, and what they are prepared to do about that language" (Dyers & Abongdia, 2010, p. 121).

EFL teachers' language attitudes may be affected by complex factors. For example, they may be affected by training and education. Baker (2011) looked at the attitudes and practices of five EFL teachers (see section 4.2.2.1 below for more details on this study). One of the teachers in her study confirmed her interest in and her focus on teaching segmentals only, particularly consonants, while the other teachers she investigated focused on both segmentals and suprasegmental in their

teachings. Baker (ibid.) attributed the teacher's focus on segmentals only to her lack of previous training and education; this particular teacher has not had previous training or education on teaching prosodic features, and thus she did not focus on teaching these features, while the other teachers had previous training and education on teaching prosodic features, and thus they focused on these features in their teaching.

Language attitudes may also be affected by age and professional experience (Agathopoulou, 2010; Arrieta, 2016; Libben & Rossman-Benjamin, 1992; Richards, 1996; Shulman, 1987). For example, Agathopoulou (2010) investigated the attitudes of prospective teachers who were university students. One of the aspects she investigated was the advantage of early instructed language learning. She found that the participants in her study agree with the significance of such learning. She further investigated the effect of age and teaching experience on her participants' attitudes. She compared the attitudes of her participants with the attitudes of older and experienced teachers and postgraduates investigated in other two studies. She found that the old and experienced participants in the other two studies were skeptical about the advantages of an early start in instructed language learning, in contrast to the participants in her study. Thus, she attributed the different attitudes adopted by the participants in her study to age and experience; the participants were young and less experienced than the other participants in the other two studies and thus they had different attitudes. Language attitudes may also be affected by experience as a language learner (Golombek, 1998; Libben & Rossman-Benjamin, 1992; Woods, 1996). For example, Woods (1996) reported a teacher whose language attitudes were affected by his learning experience. This teacher had formal instruction on French for years, but this instruction did not help him to communicate in French effectively. After his work at a company with French speakers for six months, he has developed his communications skills remarkably. As a result of this learning experience, this teacher had preference for communicative techniques in his teaching over grammar-based techniques.

In line with the factors affecting EFL teachers' language attitudes, Garret (2010) stated that language attitudes are dynamic and changeable in response to social situations. Accordingly, particular language attitudes are acquired from the social environment, such as through hearing people talking about a particular language or culture in a particular way, and by exposure to teachers' pedagogical views and choices at school (ibid.). As a result of this, EFL teachers' language attitudes may cause L2 learners, for example, to stereotype particular language varieties, and thus

this may affect their learning behaviours such as considering certain varieties but not others for learning pronunciation (McKenzie, 2008).

4.2.2 EFL Teachers' Language Attitudes Research

There is limited research into EFL teachers' language attitudes, especially in L2 pronunciation (Baker, 2014; Baker & Murphy, 2011). Until recently, there is no research investigating EFL teachers' language attitudes in different ESL and EFL contexts, especially regarding L2 pronunciation (Buss, 2016). In this regard, Kubanyiova and Feryok (2015) argued that little is known about EFL teachers' language attitudes and how these attitudes are developed, particularly regarding L2 pronunciation. The area of research concerning EFL teachers' attitudes to intelligibility and nativeness in particular is significantly lacking in investigation, especially in terms of L2 pronunciation in general and L2 pronunciation assessment in particular (Bøhn & Hansen, 2017).

The current study has looked at EFL teachers' language attitudes. Specifically, it has looked at NNS and NS EFL teachers' attitudes to intelligibility and nativeness in English pronunciation assessment in Saudi Arabia. It has investigated the attitudes through direct and indirect methods (see section 4.2.3 below for more details on these methods); it examined the attitudes of the teachers through practical tasks (indirect method) and an online questionnaire (direct method) in order to find out if the teachers' beliefs were reflected in their practices. The study adopted the LFC as a measure for intelligibility of pronunciation. It also compared between the NNS and NS EFL teachers in terms of their attitudes.

The following subsections present a review of studies looking at EFL teachers' language attitudes and have found results relevant to the current study. Interestingly, assuming all relevant studies have been found, most of them are quite recent; all the studies were conducted in 2011 and later, except for only three that were conducted in 2000, 2002 and 2005. This suggests that the research into EFL teachers' language attitudes is not extensive.

Given the limited amount of research into EFL teachers' attitudes to intelligibility and nativeness in pronunciation assessment, the studies that have considered EFL teachers' attitudes to intelligibility and nativeness in pronunciation in general or in pronunciation teaching were considered in the review below besides the studies that have looked at EFL teachers' attitudes to intelligibility and nativeness in pronunciation assessment only. Thus, in the subsections below, (a)

the majority of the studies have investigated EFL teachers' attitudes to intelligibility and nativeness in pronunciation teaching only; (b) a small number of the studies have looked at these attitudes in pronunciation assessment besides pronunciation teaching; (c) a small number of the studies have considered these attitudes in pronunciation in general; and (d) only a single study has looked at these attitudes in pronunciation assessment alone.

Finally, the studies reviewed below are divided into three types based on their findings: (1) six studies have found that EFL teachers had more orientation to nativeness than intelligibility; (2) seven studies have found more orientation to intelligibility; and (3) six studies have found roughly equal orientation to both intelligibility and nativeness. The studies of each type are presented below separately under a single subsection.

4.2.2.1 Nativeness Orientation Research

Six studies in the literature have looked at EFL teachers' attitudes to intelligibility and nativeness in L2 pronunciation and have found that the EFL teachers were more oriented to nativeness than intelligibility. For example, Tsui and Bunton (2000) found the EFL teachers in their study were more oriented to nativeness than intelligibility.

Tsui and Bunton (ibid.) investigated the attitudes of EFL teachers in Hong Kong as exhibited either implicitly or explicitly during communication via electronic messages over a period of two years. The discourses of both NNS and NS EFL teachers were analysed. The data for this study included more than a thousand language-related messages. For the two-year period from October 1997 to October 1999, the researchers collected a total of 850 messages posted by teachers. The majority of the teachers involved in posting messages (N = 85) are English NNS, compared with 17 English NS. In analysis of the data, Tsui and Bunton (ibid.) reported on many aspects of the teachers' attitudes, but only the aspects that are relevant to the current PhD study are reported here.

Tsui and Bunton (2000) found that the majority of the NNS teachers' messages explicitly or implicitly accepted NSs as a source of authority. However, these teachers were different in terms of accepting NSs as a source of authority. At one extreme, there was full acceptance of NSs as the norm. At a different level, there were cases where the EFL NNS teachers referred to NNSs as an equal source of authority as NSs. Similarly, according to NS EFL teachers, the most frequent source of authority was NS usage. They frequently considered NSs' use of linguistic forms the standard use.

Also, they frequently identified themselves as the source of the authority when English is concerned. Tsui and Bunton (ibid.) concluded that the analysis in their study showed that the NS and NNS teachers' attitudes agreed with those of the government and the business community, which showed preference for standard English (American and British Englishes) in formal communication.

Similarly, Sifakis and Sougari (2005) found that in general the EFL teachers in their study showed more orientation to nativeness than intelligibility. Sifakis and Sougari (ibid.) examined Greek EFL teachers' beliefs about English NS accents and their own English accents using a survey. The findings revealed that the teachers were generally very happy with their own accents as they believed they sounded nativelike. When the teachers were asked whether their students should have acquired a nativelike accent, the majority of the teachers believed it was very important, but some of this majority considered encouraging intelligibility in communication as well. Sifakis and Sougari (ibid.) further asked the teachers how often they used specific methodological practices. Some of the practices that were investigated and are relevant to the current PhD study were conversations among NSs and conversations among NNSs. The majority of the teachers claimed they frequently exposed their students to conversations among NSs, while only some of the teachers claimed they frequently exposed their students to conversations among NNSs. Finally, when the teachers were asked about their perspectives on the ownership of English, a small minority of the teachers said that English belonged to all speakers of English, whereas the great majority chose NSs or speakers with native competence as the owners of English. Sifakis and Sougari (ibid.) concluded that their study showed that NS norms were dominant in the Greek teachers' beliefs regarding their own pronunciation and pronunciation teaching.

Young and Walsh (2010) also found similar results like those of Sifakis and Sougari (2005) and Tsui and Bunton (2000). In a part of their study, Young and Walsh (ibid.) looked at the varieties EFL teachers taught or wanted to teach and what EFL teachers understood about ELF. The participants in their study were 26 NNS EFL teachers. They were, at the time of study, attending the same university in the UK for different degrees and subjects including a master's, MPhil or PhD degree in Applied Linguistics, Cross Cultural Communication, Education or TESOL. Their teaching experience ranged from 2 to 15 years, and their ages ranged from 23 to 40 years old, and the majority were female. The data of this study were collected using interviews only.

In the interviews, the teachers were asked about the English variety they were teaching. The majority (81%) confirmed they were teaching American English (AmE), whereas only four

teachers (15%) stated they were teaching British English (BrE), and only one teacher (4%) confirmed teaching a local variety of English (Young & Walsh, 2010). However, none of the teachers confirmed teaching an ELF variety. The teachers were then asked about the variety of English they wanted to teach. The majority (84%) referred to AmE, while only four teachers (15%) mentioned BrE, and other three teachers (12%) preferred a local variety of English. However, one teacher (4%) preferred teaching ELF. The teachers were further asked what ELF was and how useful and attractive it was. Most of the teachers (73%) expressed positive attitudes towards the concept, but only one teacher (4%) believed in it as a useful model for teaching. The final question the teachers was asked is about the variety they would teach in the future. All the teachers thought of AmE, but five of the teachers (18%) referred to BrE as well, and other four (15%) considered ELF in addition to AmE.

Young and Walsh (2010) concluded that the majority of the teachers in their study confirmed teaching standard English—American English in most cases. If given a choice, these teachers stated they would choose to teach standard English. This position was upheld by the teachers although they acknowledged that it did not correspond with the real situation of English use today around the globe. When the teachers were asked about ELF, they found it conceptually attractive, but their position was probably based on the fact they were unsure about the nature of ELF (Young & Walsh, ibid.).

Ahn (2011) also found similar results as the three studies mentioned above. In a part of her study, Ahn (ibid.) looked at the beliefs and instructional practices of one teacher in terms of native speaker norms. The teacher is a South Korean non-native speaker of English, and the context of the study is a South Korean University. The findings of this case study showed that the teacher stressed the importance of following native English norms in pronunciation, specifically the American English variety which he described as the standard. The teacher also used materials and evaluated L2 learners based on the American English variety. Ahn (ibid.) concluded, regarding the teacher' beliefs, that American English was the norm that the teacher believed it worked for L2 pronunciation learning. The teacher also reflected this belief in the instructional materials he used for teaching and in his teaching and assessment practices.

Like the studies mentioned above (i.e., Ahn, 2011; Sifakis & Sougari, 2005; Tsui & Bunton, 2000; Young & Walsh, 2010;), Baker (2011a) showed that the EFL teachers in her study were more oriented to nativeness than intelligibility. As a part of her study, Baker (ibid.) investigated EFL

teachers' beliefs regarding some linguistic features in pronunciation instruction. The researcher did not mention whether the teachers are NS or NNS. In addition to the five teachers investigated, the researcher reported her own orientation. Baker (ibid.) held semi-structured interviews with five ESL teachers from an English programme. The teaching experience of these teachers (including the researcher herself) ranged from 2 to more than 20 years. Their teaching qualifications included a TESOL certificate, MA in TESOL/Applied Linguistics, and PhD in Linguistics. Four of these teachers confirmed having formal training in teaching English pronunciation, which included training on phonology and/or sound systems of English. The percentage of teaching they devoted to pronunciation varied greatly; the lowest percentage was 10 and the highest was 80.

In the interviews, one teacher confirmed his focus on pronunciation instruction, and he mentioned his focus on things like suprasegmentals (Baker, 2011). Another three teachers reported that suprasegmentals were the focus of their instruction. Two other teachers stated their focus on both segmentals and suprasegmentals. In the interviews, the teachers were further asked on their beliefs about what pronunciation features were most important in teaching. One teacher stated that consonants were the most important. Another teacher mentioned that sentence stress was one of her priorities. Another teacher referred to intonation and rhythm. One more teacher mentioned word stress and sentence stress. Similar to this teacher, a different teacher pointed word stress out in addition to intonation and rhythm. The last teacher prioritized intonation, rhythm, and word and sentence stress. The teachers were also asked about their future needs in terms of improving their knowledge of pronunciation pedagogy. All but one of the teachers stated their need for more education on teaching prosodic features.

Baker's (2011) conclusion regarding her study did not address whether the teachers' reported practices and beliefs indicated the teachers' orientation to nativeness or intelligibility. However, considering the teachers' overall practices and beliefs, all the teachers believed in the importance of suprasegmentals, except one teacher who believed in both suprasegmentals and segmentals. Given that all suprasegmentals but sentence stress are not intelligibility features based on the LFC (see section 3.5, Chapter 3, for details on the LFC), the teachers' practices and beliefs showed more orientation to nativeness than intelligibility as defined by the LFC. The teachers' beliefs were reflected in their reported practices and in the knowledge they wanted to improve in the future.

One last study that found orientation to nativeness among EFL teachers is Stanojević et al. (2012). In a part of their study, Stanojević et al. (ibid.) looked at the attitudes of three Croatian EFL teachers as regards native English accents. To collect the data from these teachers, the researchers used an interview. In the interview, the teachers reported that they felt that native models were appropriate for teaching pronunciation; they stated that the British and American Englishes should be the models for teaching pronunciation. Further, the teachers showed some approval for the ELF pronunciation, especially when communication was concerned, but they refused to accept it as a teaching model. Stanojević et al. (ibid.) concluded that the teachers in their study were more oriented to nativeness as they preferred native models, particularly British and American Englishes, for pronunciation teaching and rejected the ELF pronunciation.

4.2.2.2 Intelligibility Orientation Research

Seven studies in the literature have found that EFL teachers were more oriented to intelligibility than nativeness regarding English pronunciation. For example, Timmis (2002) found that the EFL teachers in his study showed more orientation to intelligibility than nativeness. In a part of his study, Timmis (ibid.) used a survey questionnaire to look at teachers' views regarding whether students should have conformed to native-speaker norms of English in pronunciation. The questionnaire asked the participants to report on their beliefs about NS norms and was distributed to 180 NS and NNS teachers from 45 different countries. The findings of the questionnaire showed that most teachers, especially NS teachers, preferred intelligibility to nativeness. Timmis (ibid.) concluded that the teachers in his study seemed to be moving away from NS norms.

Like Timmis (2002), Foote et al. (2012) found that the teachers in their study were more oriented to intelligibility than nativeness. In their study, Foote et al. (ibid.) investigated the following two research questions:

1. How much, how, and with which materials is pronunciation being taught?
2. What are instructors' beliefs and attitudes toward pronunciation instruction?

Foote at al. (2012) conducted an online survey that consisted of 45 questions eliciting information about various aspects including teaching and learning materials and activities used and attitudes about pronunciation teaching. Multiple-choice questions, yes/no questions, checklists, Likert scales, and open-ended questions were all used in the survey. The participants in this study comprised 129 EFL teachers. Most of these teachers were in their 40s or 50s (63%) at the time of

the study; other 16% of the participants were over the age of 60, while other 21% were under the age of 40. The majority of the teachers are females (89%). Also, the majority are native speakers of English, besides only 28 teachers whose first languages are other than English. Almost half (48%) of the teachers had a teaching experience of 10 years or less, whereas the lengths of others' experience were 10-15 years (20%), 15-20 years (16%), or more than 20 years (16%). Almost half (49%) of the teachers had a TESL diploma from a college or university, while fewer (21%) had a master's in TESL, and further fewer (4%) had a BEd in TESL, and only one participant had a PhD in TESL.

In the survey, the EFL teachers were asked about the activities that were most useful for developing students' pronunciation (Foote at al., 2012). Segmental activities were the most considered activities by the EFL teachers, followed by suprasegmental activities, and then the activities that considered both segmental and suprasegmentals. The EFL teachers were also asked to state whether they agreed or disagreed with the statement "only native speakers should teach pronunciation". 57% of the teachers disagreed with this statement, 28% accepted it, and the remaining 15% were undecided. The teachers were also asked about the statement "there is an age-related limitation on the acquisition of native-like pronunciation". 55% expressed their agreement with this statement, 32% disagreed, and the rest (13) were undecided. They were further asked about the statement "pronunciation teaching should help make students comfortably intelligible to their listeners". The majority (89%) accepted this statement, while the rest disagreed with it. One last statement is "the goal of a pronunciation program should be to eliminate, as much as possible, foreign accents". The majority (83%) rejected the idea suggested by this statement, while 12% accepted it.

In their conclusion, Foote at al. (2012) did not mention whether the teachers in their study had orientation to nativeness or intelligibility based on their responses in the survey. However, based on the teachers' responses, it can be concluded that the majority of the teachers had more orientation to intelligibility than nativeness. This was firstly indicated by the responses to the question asking about the activities that were most useful for developing students' pronunciation; the teachers' responses to this question showed that segmental activities were the most useful, followed by suprasegmentals activities (as mentioned above). Given that most segmentals are intelligibility features and most suprasegmentals are nativeness features (according to the LFC; see section 3.5, Chapter 3), the teachers seemed to be generally more oriented to intelligibility than nativeness in this regard. The teachers' orientation to intelligibility was also indicated by their responses to the Likert-scale statements. Accordingly, the majority of the teachers agreed with the

statements that were oriented to intelligibility and disagreed with the statements oriented to nativeness.

In a PhD study based in Norway which examined teachers' attitudes towards native speaker norms in pronunciation, Sannes (2013) has reached findings similar to those of Foote at al. (2012) and Timmis (2002; mentioned above). In a part of her study, Sannes (ibid.) investigated the following question: what attitudes do EFL teachers have as to varieties of English in a classroom context?

Sannes (2013) conducted interviews with three teachers from three different schools in Norway to answer the above-mentioned question. However, she did not mention if the teachers are English NS or NNS. All these teachers considered communication as a main purpose for Norwegian students learning English. They also agreed that ELF has become more and more relevant, particularly for travelling, study and careers. Sannes (ibid.) did not discuss if the teachers were more oriented to intelligibility or nativeness. However, based on their responses, the teachers showed generally orientation to intelligibility. They showed this by placing importance to communication and ELF rather than nativeness.

Baker and Burri (2016) also found that the EFL teachers in their study had more orientation to intelligibility than nativeness. Baker and Burri (ibid.) looked at EFL teachers' perceptions as regards providing L2 learners with feedback on their pronunciation. The context of the study is an EAP programme in USA. The participants for this study were five EFL teachers with at least six years teaching experience in USA and overseas. They all had a master's degree in TESOL at the time of the study. It was not mentioned by the researchers if the participants are English NS or not. The data for the study were collected using three semi-structured interviews, five classroom observations, and two stimulated recall interviews, which are a research technique in which participants watch a video of their behaviour and then reflect on the behaviour videoed (Baker & Burri, ibid.). The three semi-structured interviews were conducted at three different times of an academic year. The interviews were used to investigate a variety of aspects including the teachers' previous and current teaching experience. Specifically, the interviews investigated the way in which the teachers taught pronunciation-related topics and the reasons behind their teaching approaches. The classroom observations and stimulated recall interviews were used to complement the data of the three semi-structured interviews. Each teacher was observed at five different times throughout one semester. The observation sessions were video-recorded, and for the pronunciation-related

activities, they were transcribed verbatim. Following observation, stimulated recall interviews with the teachers were conducted.

The findings of Baker and Burri (2016) highlighted several important themes from the interviews and observations regarding the teachers' beliefs about providing feedback on pronunciation and how they did that. However, only the themes that are relevant to the current PhD study are considered here. Accordingly, the interviews and the observations showed that feedback was essential and encouraged for improving students' intelligibility in pronunciation, while nativeness was irrelevant. All the EFL teachers stressed that the purpose of their classes was not to teach their students how to sound nativelike, but how to be intelligible. Baker and Burri (ibid.) did not discuss if the EFL teachers in their study had more orientation to intelligibility or nativeness. However, based on their findings, it can be concluded that the EFL teachers' attitudes were more oriented to intelligibility than nativeness.

Couper (2017) also found that the EFL teachers in his study show more orientation to intelligibility than nativeness (Couper, 2016, conducted a similar study but found different findings; see section 4.2.2.3 below). Couper (ibid.) explored how EFL teachers in New Zealand perceive pronunciation teaching using semi-structured interviews; they explored the perceptions of these teachers based on their cumulative knowledge and experience. The participants in this study were 19 EFL teachers with a range of ages. They were recruited from five different types of institution, including tertiary institutions, private language schools and community-focused organisations. Seventeen of the EFL teachers are NS, and further two EFL teachers are NNS. All of these teachers had a range of experience and qualifications.

The findings of the interviews in Couper (2017) showed that all the EFL teachers stated they neglected stress and intonation as they could not describe them and did not know how to teach them to their students, although they accepted they were important. However, most teachers gave segments central focus in their teaching. When the teachers were asked about their goals of pronunciation teaching, the majority stated that their goals were achieving intelligibility and effective communication. However, one of the NNS teachers expressed that pronunciation teaching should have helped make L2 English learners sound more nativelike to be intelligible. This teacher felt worried her Russian accent made her unintelligible, and thus this made her think that accentedness made speakers unintelligible. Also regarding ELF, the teachers accepted their learners' needs for ELF besides the native English norms.

Couper (2017) also reported findings in terms of identity which indicated that teachers' attitudes in his study were more oriented to intelligibility than nativeness. For example, one of the teachers expressed her concern about imposing a particular model, where she said, "I am quite nervous of being 'this is the English.' I have kind of been 'Ms. BBC' in the past and don't want to get into that" (ibid., pp. 832-833). Thus, this indicated she was moving away from a native model towards accepting any model including NNS models. Another teacher mentioned that she found some of her students were trying to sound nativelike, but she encouraged them to be proud of their own identities and helped them with their pronunciation. Similarly, another teacher told her students not to worry about their own accents and that their accents indicated who they were.

Couper (2017) did not discuss the findings of his study in light of attitudes to intelligibility and nativeness in pronunciation. However, based on his findings it can be concluded that the teachers had in general more orientation to intelligibility than nativeness based on their reported pronunciation teaching practices and beliefs. This was indicated by the majority's reported goals as they said that intelligibility and effective communication were their goals of pronunciation teaching. It was also indicated by the majority's focus on segments in pronunciation teaching and their expression of the need for ELF. The teachers' orientation to intelligibility was finally indicated by the focus of some teachers on making their students proud of their identities and own accents while opposing nativeness.

Another study with similar findings is Buckingham (2015), which looked at the perceptions of NNS and NS EFL teachers in Oman as regards native and non-native English pronunciation and their own English accent. Buckingham carried out an online survey to investigate the teachers' perceptions. The data collected included the teachers' beliefs and reported practices regarding English pronunciation.

The findings of Buckingham's (2015) study that are relevant to the current PhD study showed that the teachers did not focus on a particular accent; instead, the great majority considered "clear English" to be more important. When the teachers were asked about the variety they preferred for the didactic materials, the majority preferred either a NNS variety or NS variety to be used, while few went for a NS variety only to be used in the didactic materials. When they were asked about their own English accents, the majority were satisfied with them. Yet, a few were dissatisfied with their English accents, and they justified their dissatisfaction on two grounds: their accents did not sound nativelike, or their accents were not intelligible. The teachers were finally

asked to recommend a variety to be used in pronunciation teaching. A few teachers recommended the adoption of an ELF approach rather than particular accents, but a smaller number of the teachers recommended the use of the American and British accents as a model. Buckingham (ibid.) concluded that both the NNS and NS EFL teachers in her study generally placed more importance on teaching clear English pronunciation rather than specific accents. Also, both groups of teachers placed some importance on the global use of English.

One last study with similar findings to those above is Bøhn and Hansen (2017). Bøhn and Hansen (ibid.) conducted their study in Norway; they used an interview followed by a Likert scale questionnaire to elicit EFL teachers' attitudes regarding intelligibility and nativeness. The Likert scale questionnaire had two sections. One section with statements oriented to nativeness or intelligibility, and another section with statements concern four linguistic features: segmentals, word stress, sentence stress and intonation. The interview's participants totalled 21 EFL teachers: three NSs and the rest are NNSs. Their teaching experience ranged from one year to 32 years, with an average of 11.2 years. All but one had experience of oral English assessment. The questionnaire's respondents totalled 46 EFL teachers: two NSs and the rest are NNS. These teachers had on average 6-10 years of experience, and 32 teachers had experience of oral English assessment.

The interviews' findings showed significant variation in the interviewees' responses regarding nativeness: five (out of 21) interviewees did not find nativeness important at all, five found it slightly important, five found it somewhat important, and six found it very important (Bøhn & Hansen, 2017). The findings of the interview, however, showed that the interviewees attached significance to intelligibility. The overall findings of the Likert-scale questionnaire showed that the respondents believed that intelligibility was more important than nativeness. They also showed that most of the respondents strongly agreed on the importance of segmentals, but fewer respondents agreed on the importance of word stress and sentence stress. As for intonation, the respondents were neutral about it: they neither strongly agreed nor strongly disagreed on the importance of it. Later, four items were added to the questionnaire to verify its previous findings. The findings of these four items corroborated the questionnaire's findings with the exception of intonation, where the respondents agreed slightly more strongly on its importance this time.

Bøhn and Hansen (2017) concluded that the teachers interviewed were strongly oriented to intelligibility than nativeness, and that the teachers who responded to the Likert-scale questionnaire were strongly oriented to intelligibility than nativeness in the first section of the

questionnaire. As regards the second section of the questionnaire, Bøhn and Hansen (ibid.) concluded that:

> [T]he teachers found segmental features to be the most prominent, followed by word stress and sentence stress. Their attitudes toward intonation were somewhat less clear, partly due to poor operationalization of the construct, but the analysis suggests that they were either not as concerned with this feature or unsure of how to relate to it". (ibid., p. 65).

Based on this conclusion, the teachers who responded to the questionnaire may have been more oriented to intelligibility than nativeness in the second section of the questionnaire as far as segmental features, word stress and sentence stress were concerned. This could be indicated by the great importance given to the segmental features, which most of them are intelligibility features (according to the LFC; see section 3.5, Chapter 3), and the secondary importance given to sentence stress, which is an intelligibility feature. This is in contrast to word stress, which is a nativeness feature; word stress was given only secondary importance. However, the case of intonation was not definite. The teachers may have been oriented to intelligibility if they were concerned with intonation; otherwise, their attitudes to intonation was not obvious since they did not understand it. Overall, given the attitudes of the teachers in both the interviews and the questionnaire, the teachers were generally more oriented to intelligibility than nativeness.

4.2.2.3 Nativeness and Intelligibility Orientation Research

Six studies in the literature have looked at EFL teachers' attitudes to nativeness and intelligibility in pronunciation and found almost equal orientation to both nativeness and intelligibility. For example, Saito (2014) found almost equal orientation to both nativeness and intelligibility when he investigated EFL NNS and NS EFL teachers' perspectives regarding the characteristics of intelligible pronunciation. The participants for his study included 61 NSs and 59 native Japanese with similar teaching backgrounds in EFL classrooms in Japan. These teachers did not have any previous training on pronunciation teaching. The study took a quantitative approach where the teachers were asked to rate a list of pronunciation features with respect to their effect on intelligible pronunciation. Based on the teachers' ratings, Saito (ibid.) grouped the pronunciation features into categories reflecting general patterns. He then ranked the groups based on the total

means of rating scores, where the first group had the most effect on intelligible pronunciation according to the teachers' ratings. Below is the ranking:

1. The segmentals /l, ɹ, ð, θ, v/
2. Word/sentence stress and intonation
3. The segmentals /æ, ʌ,f/
4. Diphthongs
5. The segmentals /p, t, k, w, n, ŋ, h/

In conclusion, Saito (2014) did not mention whether the teachers in his study were oriented to either intelligibility or nativeness based on their ranking of the pronunciation features. However, based on the teachers' ratings, it seems the teachers had almost equal orientation to both nativeness and intelligibility. The pronunciation features ranked as first and second include almost the same number of features for both intelligibility and nativeness.

Arrieta (2016) also found similar results to those in Saito (2014). In a part of her master's thesis, Arrieta (ibid.) investigated the English varieties EFL teachers wanted to teach and whether the teachers' perceptions would change after their exposure to World Englishes' pronunciations. The World Englishes' pronunciations in Arrieta (ibid.) included the pronunciations of English speakers from the three circles (see section 3.2.1, Chapter 3). Arrieta (ibid.) used an online survey to collect the data for her study. The survey also included pre- and post-video questions; the post-video survey questions came after watching a four-minute video about World Englishes. The questions were a combination of Yes/No questions and Likert scale questions. The participants for the study included fourteen EFL teachers from two schools in the United States. Eleven of those teachers were 44 years old or older, while the remaining three were under this age. All but one teacher are NS. Two teachers had a TESOL certificate, one had no training on TESOL, and the rest had a master's degree in TESOL.

The findings of the pre-video Likert scale questions showed that the average score of the teachers' agreement with the importance of World Englishes training was 2.64 out of 4 points (Arrieta, 2016), where 1 representing the least agreement and 4 the most agreement. The average score for their agreement with the importance of teaching World Englishes was 2.86. The findings of the post-video Likert scale questions showed that the teachers' average scores of agreement were

2.57 for the importance of World Englishes training, and 2.71 for the importance of teaching World Englishes to their students.

The findings of the Yes/No questions and teachers' elaborations on these questions showed that the majority of teachers stated it was slight important for their students to be exposed to World Englishes' pronunciations (Arrieta, 2016). One teacher stated that this "depends on what other Englishes they will be exposed to in their future lives. Some may experience many while others may experience only a few with regularity" (Arrieta, ibid., p. 60). However, when the teachers were asked whether they considered World Englishes in their teaching, the majority (N = 12) confirmed incorporating World Englishes materials in the classroom sometimes, and they reported drawing their students' attention to differences among English varieties from time to time.

In the conclusion of her study, Arrieta (2016) did not discuss whether the teachers in her study showed more orientation to nativeness or intelligibility. However, given that World Englishes include NNS varieties of both Outer and Expanding circles (see section 3.2.1, Chapter 3), recognising the importance of World Englishes and considering them in teaching and learning could indicate accepting NNS varieties as intelligible varieties while not adhering to native varieties. This in returns indicates orientation towards intelligibility rather than nativeness. Thus, based on Arrieta's (ibid.) findings, it seems that the teachers were almost equally oriented to nativeness and intelligibility. The teachers in her study showed average agreement with the importance of World Englishes teaching and learning. The teachers also confirmed teaching World Englishes only from time to time.

Similar to the other studies in this section, the conclusions of Buss (2016) showed that the EFL teachers in her study were almost equally oriented to both intelligibility and nativeness. In a part of her study, Buss (ibid.) has explored the self-reported practices of Brazilian EFL teachers regarding pronunciation and the beliefs of these teachers regarding pronunciation learning and teaching. Buss (ibid.) collected the data for her study via an online survey. This survey included a total of 74 open- and closed-ended questions. The types of the questions included short answer, multiple choice, and Likert-type questions. A total of sixty Brazilian Portuguese-speaking EFL teachers were recruited for this study. The majority of these teachers (71%) were female, and the majority (65%) were in their 20s or 30s. All the teachers were working at private language schools or universities, while a few were working at elementary schools. Over half of the teachers had

taught English for more than ten years, and the majority had at least five-year teaching experience. The teachers had different qualifications ranging from an undergraduate degree to a PhD.

Firstly, the teachers in Buss's (2016) study were asked about the percentage of class time they spent on teaching suprasegmental features. The majority of teachers (60%) stated they spent between 0% and 40% of the class time teaching suprasegmentals. Slightly over 20% of the teachers stated they spent a half of the class time teaching suprasegmentals, and only 18% of the teachers told they spent 60% or more of the class time on suprasegmentals. The teachers were further asked how frequently they taught specific features of pronunciation. The sounds such as /θ/ and /ð/ were the most frequently taught features; 89% of the teachers teach these sounds all the time or often. The second most frequently taught features were suffixes such as *-ed* and *-s* endings; 86% of the teachers taught them always or often. Word stress came next with 72% of the teachers teaching it frequently, then followed intonation (55%). Overall, the findings indicated that word-level features (e.g., word stress and intonation) appeared to be more highlighted by the teachers than the other pronunciation features.

The teachers were further asked how L2 learners learn pronunciation (Buss, 2016). The findings showed that importance of native models was a minor topic within the teachers' comments and recurred for only seven times through the comments. The teachers were also asked whether they agreed or disagreed with different statements about pronunciation teaching oriented to intelligibility or nativeness. The responses to these statements showed that the teachers were generally oriented to intelligibility. For example, the responses to one statement that concerned the objective of teaching pronunciation, a great majority of the teachers agreed that it should have assisted the students to be easily intelligible to their listeners. Also, a great majority of the teachers disagreed with another statement saying that the goal of pronunciation teaching is to eliminate, as much as possible, a foreign accent. The majority also agreed with a third statement saying that a NS EFL teacher is not necessarily the best to teach pronunciation. Finally, almost half of the teachers agreed with a fourth statement saying that native pronunciation is not required as a model for pronunciation teaching.

Overall, the teachers' attitudes in Buss's (2016) study were almost equally oriented to intelligibility and nativeness. The teachers' reported beliefs in Buss's (ibid.) study showed more orientation to intelligibility than nativeness. Buss (ibid.) mentioned that the majority of the teachers in her study stated that "intelligibility and comprehensibility are better goals for pronunciation

teaching than foreign accent reduction" (p. 633). Buss (ibid.) further mentioned that "there is a tendency [, among the teachers in her study,] to think that native speech does not even need to be the model for pronunciation instruction" (p. 633). However, the teachers' reported practices in Buss's (ibid.) study showed more orientation to nativeness. Buss (ibid.) did not discuss whether the teachers in her study were more oriented to intelligibility or nativeness based on their reported practices, but it seems they were more oriented to nativeness in this regard since they reported their focus on phonological and phonetic features that are nativeness features (according to the LFC), such as dental fricatives.

Another study that found EFL teachers were almost equally oriented to both intelligibility and nativeness is Couper (2016) study (Couper, 2017, found different results; see section 4.2.2.2 above). In this study, Couper (ibid.) investigated EFL teachers' concerns and issues as regards pronunciation teaching in Uruguay through individual semi-structured interviews. The participants in this study were 28 EFL teachers: 25 females and three males. They had a range of qualifications and teaching experience; the shortest teaching experience was five years. However, Couper (ibid.) did not mention if the EFL teachers in his study are NS or NNS.

When talking about the importance of teaching pronunciation in the interviews, one of the teachers stated that NNS teachers sometimes did not teach pronunciation since they did not feel confident to teach it as they are NNS. When talking about their activities and techniques in the classroom, one of the teachers reported his use of NS recordings for teaching listening so that the students could get prepared to interact fluently. This teacher further said that exposing students to NS speech is the ideal way to help them acquire pronunciation, but he still thought that it was important to expose the students to both NNS and NS varieties of English. Another two teachers reported their focus on exposing their students to both NNS and NS varieties of English as they believed it was important. However, for some other teachers, the traditional NS model was their focus when teaching pronunciation. For example, one teacher reported the need to reduce accentedness in his students. However, although nativeness was the focus for those teachers, most of them stated the need to focus on intelligibility as well.

Couper (2016) also investigated the teachers' practices in teaching suprasegmentals and segmentals. Some of the teachers reported their use of stress and intonation in teaching listening where they asked their students to guess the context and situation using these two features (Couper, ibid., did not give an example of this). Some others reported using minimal pairs in listen-

and-repeat practices for differentiating between similar sounds; one of these teachers gave the short and long vowels in the words "live" and "leave" as examples. Two teachers talked about their experience with trying to develop their students' awareness and pronunciation of regular past tense morpheme *-ed* and sounds such as the final *-s* in third person singular and in the plural. As regards intonation specifically, one teacher reported drawing her students' attention to the importance of intonation and asked them to focus on it. Another teacher reported his use of drama to teach intonation, where he had his students read a play out loud and use intonation to highlight different moods and emotions.

Couper (2016) did not discuss his findings in terms of the teachers' orientation to nativeness or intelligibility. However, his study showed in general the teachers' orientation to both intelligibility and nativeness in terms of English pronunciation teaching. The study reported that some teachers confirmed their teaching of phonetic and phonological features which are intelligibility features (according to the LFC) such as the difference between the short and long vowels, while some other teachers confirmed their teaching of phonetic or phonological which are nativeness features such as intonation.

One more study with similar findings is Khatib and Monfared (2017), which looked at NNS and NS teachers' beliefs about the significance of NS accents, their preferences in relation to pronunciation norms, and ownership of English. The participants for this study totalled 352 EFL teachers: 112 NSs and 240 NNSs. The teachers are both males and females from different age groups and with a range of educational backgrounds and teaching experience. The data for this study were elicited using a survey questionnaire containing both close-ended and open-ended questions. Interviews were also conducted with 20 NS EFL teachers and 40 NNS EFL teachers who had formerly completely the questionnaires.

Firstly, the teachers were asked in the questionnaire whether it was more important to have nativelike pronunciation or to attain clear and intelligible pronunciation in communication (Khatib & Monfared, 2017). The responses to this question indicated that the majority of the NNS and NS teachers believed that intelligible pronunciation was more important. When the teachers were interviewed, they showed similar attitudes in this regard. The NNS EFL teachers, apart from the NS teachers, were also asked about their satisfaction with their accents. The NNS EFL teachers claimed to be generally happy with their own accents. Teachers were also asked about the most appropriate pronunciation accent for their students in the classroom. The great majority of the NNS

EFL teachers preferred native pronunciation, but the NS teachers placed more importance on the clarity and intelligibility of pronunciation rather than native accents. The teachers were further asked about their views regarding the accents of their students. The majority of the NNS and NS EFL teachers held that NS accents were not relevant as long as communication was not adversely affected. When asked about their views regarding NNS teachers' accents, almost a half of the NNS EFL teachers expressed the need for EFL teachers to have a nativelike accent in contrast to the other half of teachers who prioritized NNS accents with more focus on meaning. However, the majority of the NS EFL teachers promoted NNS accents and highlighted the need for intelligibility of pronunciation.

Khatib and Monfared (2017) further asked the teachers in their study about using materials with conversations between NSs and between NSs and NNSs in the classroom. Almost half of the NNS and NS EFL teachers claimed to include conversations between NSs in the classroom regularly, and the other half of teachers claimed to use conversations between NSs and NNSs. In one last question, the teachers were asked about the ownership of English. Almost half of the NNS and NS EFL teachers stated that English belonged to NSs, while the other half stated that English belonged to no one.

It was concluded from this study that a half of the NNS EFL teachers were "prejudiced" against NNS varieties of English, and they showed preference for native pronunciation (Khatib & Monfared, 2017, p. 230). However, the other half "highly valued" their NNS English accent, but they slightly preferred native pronunciation (ibid., p. 230). However, the authors did not discuss the attitudes of the NS EFL teachers, but the findings of their study showed that the NSs were more oriented to intelligibility than nativeness across the questions whose responses were reported in the study (some responses by the NS teachers were not mentioned, and the authors did not give a reason for that).

In a further study, Monfared and Khatib (2018) found similar results. In this study they looked at the attitudes of NNS EFL teachers towards different English varieties, including NS and their own varieties. The participants for this study are 260 NNS EFL teachers from India and Iran. The Indian teachers were teaching in British Council centres in New Delhi and Mumbai, and the Iranian teachers were teaching in three English centres in Tehran. Both groups comprise a mixture of males and females with different educational backgrounds and teaching experience; their qualifications were either a BA, MA or PhD degree, and their teaching experience ranged from one

year to over ten years. The data for this study were collected using a Likert scale questionnaire. The questionnaire included 12 items: some items were concerning the cognitive aspect of attitudes, some were concerning the affective aspect and some others were concerning the behavioural aspect.

The findings on the three aspects of attitude showed that the majority of the teachers had a more positive attitude towards the intelligibility of their English accent, were significantly satisfied with their own variety of English and wanted to keep it (Monfared & Khatib, 2018). When the teachers were asked whether they preferred that their students sounded nativelike or could convey the message intended from a conversation effectively, most of the teachers preferred that their students sounded nativelike. Monfared and Khatib (ibid., p. 69) concluded that half of the teachers in their study were "prejudiced against their own variety of English and indicated their preference for native American English pronunciation". However, the other half of the teachers "highly valued" their own variety of English (ibid., p. 69).

Unlike the current PhD study, none of the studies mentioned above adopted the LFC as a framework for intelligibility, nor looked at the attitudes in Saudi Arabia. Also, unlike this PhD study, the majority of the studies reviewed above did not compare NNS and NS EFL teachers' attitudes, nor investigated EFL teachers' attitudes through indirect methods (e.g., practical assessment tasks, like the case in the current PhD study), nor investigated EFL teachers' attitudes in pronunciation assessment.

4.2.3 Popular Methodologies in EFL Teachers' Language Attitudes Research

The approaches to studying language attitudes in general are commonly divided into three major categories: direct measures, indirect measures, and societal treatment measures (Garret, 2010; Ryan et al., 1982). The most common indirect methods are the matched guise technique (MGT) and the verbal guise technique (VGT) (Litzenberg, 2013). The MGT and VGT are supposed to be the most effective methods for measuring participants' perceptions that are often hidden behind a social facade (Dalton-Puffer et al., 1995; Giles & Billing, 2004). According to Ahn (2017, as cited in Monfared & Khatib, 2018), investigating language attitudes is not an easy task since attitudes are not directly noticeable, but they can only be deduced from observable responses.

The MGT and VGT involve playing recorded speech samples of different accents to the participants and having them rate the speakers on different traits (Jindapitak, 2015). The MGT

involves the use of a single speaker producing the same script with two or more accents, dialects or languages and a listener is asked to rate the different accents with respect to various personality traits (ibid.). The VGT, on the other hand, involves the use of different speakers, who are supposedly able to present naturalistic samples of their language varieties better than a single bilingual speaker. Traditionally, language attitudes research was undertaken using the MGT and VGT, often to investigate the affective part of attitudes with regard to stereotypical impressions (Agheyisi & Fishman, 1970).

Direct methods, however, are used to investigate various kinds of language attitudes, such as how respondents view a variety of language, their language preferences, the desirability of a variety, their reasons for learning a language, and self-reporting on language use (Ryan et al., 1982). Direct methods involve the use of direct questions or interviews (ibid.). As for societal treatment measures, they include, but are not limited to, ethnographies and analysis of language policies or how languages are treated in newspapers (Garret, 2010; Ryan et al., 1982). As mentioned above, the majority of research reviewed in the sections above have not investigated EFL teachers' attitudes through indirect methods (e.g., practical assessment tasks, like the case in the current PhD study), but they have investigated EFL teachers' attitudes through direct methods (e.g., direct questions in questionnaires and/or interviews).

4.3 Summary of Chapter and Research Questions

This chapter has presented an overview on EFL teachers' language attitudes; it has defined EFL teachers' language attitudes, and it discussed the factors affecting these attitudes. It has then discussed the popular methodologies used in the literature to research EFL teachers' language attitudes. Finally, the chapter has presented a review for the research looking at EFL teachers' attitudes to intelligibility and nativeness in pronunciation.

The research into EFL teachers' attitudes to intelligibility and nativeness reviewed in this chapter have provided three main different and, to some extent, conflicting findings: six studies have shown that EFL teachers had more orientation to nativeness than intelligibility; seven studies have shown that EFL teachers had more orientation to intelligibility than nativeness; and six other studies have found almost equal orientation to both intelligibility and nativeness among EFL teachers.

Unlike the current PhD study, none of the research into EFL teachers' attitudes to intelligibility and nativeness reviewed in this chapter has adopted the LFC as a framework for intelligibility. This PhD study adopts the LFC as a framework for pronunciation assessment and briefly discusses the success of this approach in the discussion chapter (Chapter 6). Also, unlike the current study, none of the research into EFL teachers' attitudes to intelligibility and nativeness reviewed in this chapter has investigated EFL teachers' attitudes in a Saudi context. Further, most of this research has not compared NNS and NS EFL teachers in terms of their attitudes to intelligibility and nativeness, nor has it investigated EFL attitudes through indirect methods such as practical assessment tasks, nor investigated the attitudes in pronunciation assessment, unlike the current PhD study.

Given the relevant research reviewed in this chapter and the gaps of such research highlighted above, the research question set for this exploratory PhD study is:

> What evidence is there that the attitudes of the NS and NNS EFL teachers in Saudi Arabia are oriented towards nativeness or intelligibility, as defined by Jenkins' (2000) LFC, when assessing pronunciation?
>
> This main research question has the following research sub-questions:

a. Are the NS and NNS EFL teachers oriented to intelligibility or nativeness when doing an assessment of students' speech task? Are the two groups of teachers different in this regard?
b. Are the NS and NNS EFL teachers oriented to intelligibility or nativeness when doing an acceptability task? Are the two groups of teachers different in this regard?
c. Do they demonstrate attitudes towards intelligibility or nativeness in an online questionnaire? Are the two groups of teachers different in this regard?
d. Are their attitudes in the practical tasks and the online questionnaire in harmony or in conflict? If not, why?
e. Are their attitudes in the practical tasks and the online questionnaire affected by any factors such as any personal background information (e.g., age and experience)? Or such as the institutional assessment policies?

CHAPTER 5: METHODOLOGY

5.1 Introduction

This chapter describes the methodology used to address the research question of this study. It begins with presenting the research question that has led to this exploratory study. Following this, the research design for this study is explained, where the research methods, data collection instruments and sampling are explained. Ethical consideration and the participants of this study are further considered. Following this, the instruments for collecting the data of this study are described in detail along with the procedures followed in collection of the data. The process of piloting each instrument is further explained along with the amendments that were introduced to the instruments following piloting them. Finally, the reliability of the study instruments along with the analysis of the data are discussed.

5.2 Research Questions

The research question set for this exploratory PhD study is:

What evidence is there that the attitudes of the NS and NNS EFL teachers in Saudi Arabia are oriented towards nativeness or intelligibility when assessing pronunciation?

This main research question has the following research sub-questions:

a. Are the NS and NNS EFL teachers oriented to intelligibility or nativeness when doing an assessment of students' speech task? Are the two groups of teachers different in this regard?
b. Are the NS and NNS EFL teachers oriented to intelligibility or nativeness when doing an acceptability task? Are the two groups of teachers different in this regard?
c. Do they demonstrate attitudes towards intelligibility or nativeness in an online questionnaire? Are the two groups of teachers different in this regard?
d. Are their attitudes in the practical tasks and the online questionnaire in harmony or in conflict? If in conflict, why?
e. Are their attitudes in the practical tasks and the online questionnaire affected by any factors such as any personal background information (e.g., age and experience)? Or such as the institutional assessment policies?

5.3 Research Design

This PhD study adopted a mixed-methods design; it made use of both quantitative and qualitative methods, and direct and indirect methods to answer the research questions. Mixed methods have been proven to optimise validity of data collection and reliability of the findings. According to Dörnyei (2007, p. 163), "[t]here is ample evidence in the literature that combining methods can open up fruitful new avenues for research in the social sciences". For example, mixed methods can assure the trustworthiness of a study (Lincoln & Guba, 1985); so, if there are different research methods used to answer the research question of a study, the findings of such a study tend to be more reliable than a study with a single research method.

Another advantage of using mixed methods is that they achieve a more complete understanding of a target phenomenon and enable a researcher to check one set of findings against the other (Sandelowski, 2003). This is particularly clear in case of the questionnaires that are followed by practical exercises or interviews where particular target phenomena that emerge in the questionnaire can be traced in the practical exercises and the follow-up interviews, or vice versa, and the findings of a particular method can be compared with the findings on other methods. For this PhD study, there were two practical tasks (i.e., an assessment of students' speech task and an acceptability task) which were used separately to investigate EFL teachers' attitudes to intelligibility and nativeness in pronunciation assessment, followed by an online questionnaire investigating the same attitudes to further find out if the attitudes in the online questionnaire were reflected in the practical tasks. The practical tasks and the online questionnaire were followed by an interview to further investigate the attitudes found in the study tasks (i.e., the practical tasks and the online questionnaire).

One last advantage of using mixed methods is that they can inform the development of the different research methods used in the study–e.g., a questionnaire can be used to select participants for a follow-up interview study (Greene et al., 1989). In case of this PhD study, the online questionnaire and the practical tasks were used to select the participants of the follow-up interview. The participants who had behaved differently in the questionnaire and the practical exercises were drawn for a follow-up interview to ask them further about their behaviour in the study tasks (see Table 5.1 below for the type of data and instruments along with the data analysis for the sub-questions of the main research question).

Table 5.1: The type of data and instruments along with the data analysis for the sub-questions of the main research question

Sub-question	Data obtained type	Instrument	Data analysis
First part of sub-question a	Quantitative	Assessment of students' speech	Percentages were calculated for the data of the assessment of students' speech task and then two comparisons were made using SPSS between (1) the percentages calculated for NNS teachers on the intelligibility and nativeness items and (2) the percentages calculated for NS teachers on the intelligibility and nativeness items
Second part of sub-question a	Quantitative	Assessment of students' speech	Two comparisons were made using SPSS between (1) the percentages calculated for both groups of teachers on intelligibility items and (2) between the percentages calculated for both groups of teachers on nativeness items
First part of sub-question b	Quantitative	Acceptability task	Two comparisons for the data of the acceptability task were made using SPSS between (1) the scores of NNS teachers on the intelligibility and nativeness items and (2) between the scores of NS teachers on the intelligibility and nativeness items
Second part of sub-question b	Quantitative	Acceptability task	Two comparisons for the data of the acceptability task were made using SPSS between (1) the scores of both groups of teachers on intelligibility items and (2) between the scores of both groups of teachers on nativeness items
First part of sub-question c	Quantitative	Online questionnaire	Two comparisons for the data of the online questionnaire were made using SPSS: (1) between the scores of NNS teachers on the intelligibility and nativeness items; and (2) between the scores of NS teachers on the intelligibility and nativeness items;
Second part of sub-question c	Quantitative	Online questionnaire	Two comparisons for the data of the online questionnaire were made using SPSS: (1) between the scores of both groups of teachers on intelligibility items; and (2) between the scores of both groups of teachers on nativeness items
First part of sub-question d	Quantitative	n/a	Four correlation analyses for the data of both the practical tasks and the online questionnaire were made for both groups of teachers using SPSS: correlation analyses for the scores of NNS EFL teachers on intelligibility and nativeness items in the two practical tasks and the online questionnaire, and correlation analyses for the scores of NS EFL teachers on intelligibility and nativeness items in the two practical tasks and the online questionnaire
Second part of sub-question d	Qualitative	Follow-up interview	Deductive analysis for the data of the follow-up interviews was conducted. In this process, the second parts of the third and fourth sub-questions of the main research question of this study were determined as categories, and then any connections in the data of the follow-up interviews were mapped to these categories.
First part of sub-question e	Quantitative	n/a	Six correlation analyses for the data of the study tasks and the three personal background factors (i.e., age, teaching experience and speaking assessment experience) were made for both groups of teachers using SPSS: correlation analyses for the scores of the NNS teachers on intelligibility and nativeness items in the study tasks and the three personal background factors, and correlation analyses for the scores of the NS teachers on intelligibility and nativeness items in the study tasks and the three personal background factors
Second part of sub-question e		Follow-up interview	(same as the second part of sub-question c)

5.4 Instruments

As shown in Table 5.1 above, four research instruments were designed for collecting the data of this PhD study, where each task was used separately to measure the attitudes of the EFL teachers: an assessment of students' speech task, an acceptability task, an online questionnaire, and a follow-up interview. All these instruments except the follow-up interview had been piloted in advance of data collection to ensure their validity, and thus changes have been introduced to the instruments accordingly.

5.4.1 Assessment of Students' Speech Task

The assessment of students' speech task was used in this PhD study to answer the first sub-question of the main research question (RQ1a): Are the NNS and NS EFL teachers in this study oriented to intelligibility or nativeness when doing an assessment of students' speech task? Are the two groups of teachers different in this regard? The assessment of students' speech task was developed because, as far as the author of this study is aware, there are no suitable similar instruments in previous research that could be adopted or adapted to measure the attitudes of EFL teachers towards intelligibility and nativeness through their assessment of the speech of learners, particularly using the LFC as a framework for intelligibility. Following, to a small extent, the verbal guise technique (VGT; see section 4.2.3, Chapter 4, for details on the VGT), the author of this PhD study developed this task. As mentioned in section 4.2.3 (Chapter 4), the VGT involves playing recorded speech samples performed by different speakers with different accents to listeners and having these listeners rate the speakers on different personal traits such as intelligence, reliability, kindness, and so on (Jindapitak, 2015). Similarly, the assessment of students' speech task involves playing six recorded speech samples performed by three speakers, but, differently, the speakers had different proficiency levels in English and they had relatively similar accents as they are all Saudi Arabic-speaking students. Also different to the VGT, in the assessment of students' speech task, the listeners rate the speakers using a list of the core and non-core features of the LFC. The core features of the LFC are the features important for intelligibility in ELF speech and thus they are intelligibility features. The non-core features are not important for intelligibility in ELF speech and thus they are nativeness features (see section 3.5, Chapter 3). The listeners rate the speakers by checking if the speakers made mistakes with the list of both kinds of features (see section 5.4.1.1 below for more details).

5.4.1.1 The Design of Assessment of Students' Speech Task

In the assessment of students' speech task, the participants (the EFL teachers who participated in this study) assess Arabic speaking learners of L2 English by identifying the phonological and phonetic features in recorded speech samples produced by these learners. To identify the features, the participants were given a list of phonological and phonetic features to check against the recordings. The features are commonly produced by learners of English from another L1 background (Jenkins, 2000; see next paragraph). An example of these features is "deletion/insertion of sounds in initial consonant clusters (two or more consonants together)" (see Appendix I for the complete list). The participants listened to the recordings of the learners and then checked whether these learners made one or more of these phonological and phonetic features. If the participants found that the learners had made mistakes with one or more of the features in their speech, then the participants gave an example for each feature they found problematic; they gave at least an example for the feature from the learners' speech. For example, if a participant found that the phonological feature mentioned above was problematic in a learner's speech: i.e., "deletion/insertion of sounds in initial consonant clusters", the participant gave an example of it with the word which included it from the learner's speech. Examples of this feature with the words including it from the teachers' responses were "spoons" and "slabs". The reason for asking the participants to give examples of the features they found problematic rather than identifying the features only (e.g., by ticking them) was to ensure that the participants understood the phonological and phonetic features by providing examples matching them.

The list of the phonological and phonetic features, which was used in the assessment of students' speech task, was based on the core and non-core features of the LFC (see section 3.5, Chapter 3; see also Table 5.2 below). The items based on the core features of the LFC are intelligibility features, but the items based on the non-core features of the LFC are nativeness features. It was hypothesised in this PhD study that if the participants were more oriented to intelligibility, they would tend to focus more on the intelligibility features and thus they would identify more intelligibility features as problematic than nativeness features, and vice versa: i.e., if the participants were more oriented to nativeness, they would tend to focus more on the nativeness features and thus they would identify more nativeness features as problematic than intelligibility items.

Table 5.2: The core and non-core features of the LFC

The Core Features of the LFC	The non-Core Features of the LFC
Initial consonant clusters	Medial consonant clusters
The retroflex approximant [ɻ]	Final consonant clusters
The aspirated /t/	Θ ð
The aspirated /p/ (e.g., peas)	The dark /l/
The aspirated /k/ (e.g., call)	The sound [ɾ] (the American pronunciation of /t/ when it occurs intervocalically (e.g., matter))
The lengths of vowels preceding final voiced and voiceless consonants (e.g., *back* vs *bag*)	The American pronunciation of the intervocalic two sounds /nt/ when they occur before an unstressed syllable (e.g., twenty)
Shortness and length of vowels	Intonation
Shortness and length of vowels in diphthongs	
The sound /ɜ:/ (e.g., RP pronunciation of the underlined sounds in *purple*)	
The phrase-level stress (or nuclear stress)	

The list of the phonological and phonetic features included eighteen items: ten of these items were based on the intelligibility features (i.e., the features based on the core features of the LFC and which are important for intelligibility in ELF speech, as mentioned in the preceding paragraph) and the remaining eight items were based on the nativeness features (i.e., the features based on the non-core features of the LFC and which are not important for intelligibility in ELF speech, as mentioned in the preceding paragraph). However, one core feature of the LFC were not included: quality of the vowels and diphthongs. It was difficult to represent this feature in a list since identifying quality as problematic cannot alone indicate that a participant is oriented to intelligibility or nativeness regarding this feature, but further investigation is required to decide the attitudes of the participants. If the speaker is using nativelike quality but s/he is not consistent in using one particular quality, s/he may be oriented to nativeness, but if the speaker is consistent in using one particular quality, even though it is a non-native quality, s/he may be oriented to intelligibility (see section 3.5, Chapter 3). However, it was hoped that the participants were still able to comment on this feature in a section following the list of the features, where they were asked to

list any other phonological or phonetic features they found problematic in the speakers' speech (see the next paragraph). Also, this was only one feature, and it may have not counted much alone, given there were still 18 items that could inform the findings, in addition to the other phonological and phonetic features identified in the section following the list.

Following the list of the phonological and phonetic features, the participants were also given a further section in the assessment of students' speech task where they could write down any additional phonological or phonetic features they identified in the learners' speech and gave examples to these features. The significance of such a section was demonstrated by Dornyei (2007); Dornyei (ibid.) stated that "[o]pen-format items can provide a far greater richness than fully quantitative data. The open responses can […] lead us to identify issues not previously anticipated" (p. 107; see Appendices J and K for a copy of the assessment of students' speech task and the passages used in recordings).

5.4.1.2 The Learners Assessed in the Task

The learners whose pronunciations were assessed in the task were three students at KAU. They are all male, and they were undergraduate students at the time of data collection. They were doing English language courses at KAU where samples of their speech were collected for the purpose of assessment in this study (see section 5.4.1.2 below). They were advanced, intermediate and beginner learners of L2 English respectively as classified by their EFL teacher, who collected the speech samples from these students on behalf of the author (see section 5.4.1.3 below for more details). In the view of the author of this study, having L2 English learners with different levels of proficiency may have helped produce recording samples with different intelligibility and nativeness features (i.e., the core and non-core features of the LFC; see section 5.4.1.1 above), but which jointly represented all features.

5.4.1.3 The Recordings of Students' Speech Task

The recordings for the assessment of students' speech task included two kinds of speech performed by the three learners. The first kind involved reading a passage. The reading passage is one used in the *Speech Accent Archive* (Weinberger, 2013; see Appendix J), which is free for use as stated by the archive website policy. According to the methodology of the archive, the reading passage uses common English words but with a variety of difficult English sounds and sound sequences, and it contains almost all the sounds of English plus most clusters of standard American

English. However, some words were inserted by the author of this PhD study into the passage to make sure the passage incorporates all the intelligibility and nativeness features considered in this research (i.e., the core and non-core features of the LFC). The second kind involved free speech. For this kind, the learners were asked to speak for two minutes about the last event they had participated in. The learners were asked to choose whatever important event they participated in recently and they felt they could speak about it for two minutes. As a result, the total of the recordings was six speech samples: three learners orally reading a passage and speaking about an event they participated in.

To collect the recordings of the assessment of students' speech task, the EFL teacher, who is a colleague of the author of this study, was given the reading passage and the speaking task and asked to collect the speech samples from three of his students with different proficiency levels in English: advanced, intermediate and beginner learners, based on the teacher's familiarity with these learners. The teacher had the learners perform the tasks individually in a quiet classroom. The teacher recorded the students using his phone, given the lack of recording facilities on campus. However, the quality of the recordings was supposed to be adequate for the task since there was no intention to perform acoustic analysis on the samples. Finally, the recordings were collected from the teacher using WhatsApp.

5.4.1.4 Preparation for Assessment of Students' Speech Task

Prior to undertaking any assessment of the learners, the participants were presented with the scripts of the reading passage and the free speech performed by the learners so that they could familiarise themselves with the scripts. It was hypothesised that this process could help avoid the facilitation of comprehension that may have resulted from hearing the recording scripts for the first time. This facilitation of comprehension could affect assessment when listening to the recorded scripts again (albeit from a different speaker, e.g., a participant may have considered something wrong when hearing it for the first time but not for the second time).

In order to avoid any practice effect, a counterbalanced design was used in the order of the speech samples the participants rated. This means one third of the participants started assessing the beginner learner, a third started assessing the intermediate learner, and the remaining third started assessing the advanced learner.

5.4.2 Personal Background Information Questionnaire

After the assessment of students' speech task had been completed, a questionnaire was used to seek some personal background information about the participants. It sought for details on the participants' ages, first and second languages, and qualifications for EFL teaching and EFL speaking assessment. It also sought for information about the participants' EFL teaching experience and speaking assessment experience and their knowledge of the LFC. The questionnaire was used to find out about the participants' experience and knowledge, which may have had an impact on the participants' language attitudes. Correlation analyses were used in this PhD study to investigate such an impact (see section 5.12.7 below for more details on these correlation analyses; see Appendix L for a copy of the personal background information questionnaire).

5.4.3 Acceptability Task

An acceptability task was also used in this PhD study to further investigate the participants' attitudes to intelligibility and nativeness in pronunciation assessment. The acceptability task was used to answer the second sub-question of the main research question (RQ1b): Are the NNS and NS EFL teachers in this study oriented to intelligibility or nativeness when doing an acceptability task? Are the two groups of teachers different in this regard? This task was conducted using a computer software called DMDX, which is a "display system used in psychological laboratories around the world to measure reaction times to visual and auditory stimuli" (Forester, 2002, 1st para.). This display system operates using a special script written to run the visual and auditory stimuli on the system and using audio recordings that come with the script. The script uses special symbols that represent different formats on the system. It includes symbols that are responsible for many formats in the visual stimuli on the system, such as, but not limited to, controlling the layout and duration of the display, controlling the continuous or automatic running of the display, selecting the input device used for inserting responses into the display system (e.g., a keyboard), and selecting the size of the text on the display system. The script for the visual stimuli also includes any written instructions or details for the respondents, who react to the visual and auditory stimuli. The script for the auditory stimuli includes several formats including, but not limited to, the order of the audio recordings according to the appearance of their transcripts on the display system, and when the audio recordings should play (e.g., during the display of their transcripts). Examples of such scripts are available on the display system's website (http://www.u.arizona.edu/~kforster/dmdx/dmdx.htm).

The script for both the visual and auditory stimuli needs to be written on a Revisable Form Text (RFT) file, while the audio recordings need to be in MP3 format. Both the script and audio recordings files need to be saved on a single file, so that the DMDX display system can run the script along with the audio recordings. When the script is needed to be run on the DMDX display system, there is an uploading icon on the DMDX display system from where the script document needs to be selected from the computer and then run on the system. After running the whole script along with the audio recordings, the DMDX display system presents three results in a text format so that they can be copied into a document and saved for analysis. The three results are the preparation and reaction times in milliseconds along with the error rate in a percentage form (i.e., out of 100%; see section 5.12.2 below for more explanation on the error rate).

To construct the acceptability task, a list of 66 words with their International Phonetic Alphabet (IPA) transcriptions were prepared for an English NS expert in phonetics to pronounce them and her pronunciations were recorded in a soundproof room (see Appendix M for the list of words). The IPA transcriptions of the words represented (a) 22 intelligible pronunciations (for 22 words), where they included an intelligibility feature (e.g., /t̪ɪk/ for the word "thick"; see section 5.4.1.1 above for details on the intelligibility and nativeness features), (b) 22 nativelike pronunciations, where they included a nativeness feature (e.g., /θʌm/ for the word "thumb"), and (c) 22 unacceptable pronunciations (e.g., /bɪn/ for the word "thin"). It was hypothesised in this PhD study that a participant would accept more intelligible or nativelike pronunciations according to their attitudes: if a participant was more oriented to intelligibility, s/he would accept more intelligible pronunciations than nativelike ones, and vice versa: i.e., if s/he was more oriented to nativeness, s/he would accept more nativelike pronunciations than intelligible ones.

The unacceptable pronunciations were used in the acceptability task for two purposes: as a distractor in order that the participants could not guess the purpose of the study, and to ensure that the participants were taking the task seriously. The totally wrong pronunciations were developed by the author of this study, and thus their unintelligibility is totally the author's personal judgement, which might be open to question. To develop these pronunciations, the common substitutions of sounds that are allowed by the LFC were avoided, and variations between the wrong pronunciations and the correct ones in terms of place and manner of articulation were considered. For example, according to the LFC, /θ/ and /ð/ are commonly substituted by a restricted set of alternatives — i.e., /t/ and /d/, /s/ and /z/, or less commonly /f/ and /v/ (see section 3.5, Chapter 3). Thus, the totally wrong pronunciation of the word "thin" was /bɪn/, where the sound /θ/ was replaced by the sound

/b/, which is not a common substitution of /θ/ such as /t/, /s/ or /f/ (according to the LFC). Also, the totally wrong pronunciation of the word "nothing" was /nʌmɪŋ/, where the sound /ð/ was replaced by the sound /m/, which is not a common substitution of /ð/ such as /d/, /z/ or /v/. Also, the replacements /b/ and /m/ are different from the correct sounds /θ/ and /ð/ in terms of the place and manner of articulation; /θ/ is fricative dental while /b/ is stop bilabial, and /ð/ is fricative dental while /m/ is nasal bilabial. In terms of vowels and diphthongs, a place of articulation was considered for variations between the wrong pronunciations and the correct ones. For example, the wrong pronunciation of the word "beer" which is pronounced as /bɪr/ in North American English (NAmE) is /bʊ[ɻ]/, where the vowel /ɪ/ is a front vowel and /ʊ/ is a back vowel. Alternatively, the wrong pronunciations of vowels or diphthongs were developed considering the shape of mouth. For example, the wrong pronunciation of the word "wear" which is pronounced as /we[ɻ]/ in NAmE is /wæ[ɻ]/, where the vowel /e/ is pronounced with half-open mouth while /æ/ is pronounced with fully open mouth.

An intelligible pronunciation in the acceptability task could sometimes be British English (BrE) or NAmE pronunciation based on Oxford Learner's Dictionaries (available at https://www.oxfordlearnersdictionaries.com/) besides being intelligible (according to the LFC). For example, the NAmE pronunciation of /r/, which is the retroflex approximant [ɻ], is opted for in the LFC, and thus it is both an intelligible pronunciation and a NAmE pronunciation. Another example is the BrE pronunciation of /t/ in contrast to the NAmE pronunciation once it occurs intervocalically, where it becomes the voiced flap [ɾ] in the NAmE pronunciation. The BrE pronunciation of /t/ in this case is opted for in the LFC, and thus it is both an intelligible pronunciation and a BrE pronunciation. However, a nativelike pronunciation could be either a BrE and intelligible pronunciation, a NAmE and intelligible pronunciation, or both a BrE and NAmE pronunciation besides an intelligible pronunciation. For example, the sound /t/ is an intelligible pronunciation of the sound /θ/ according to the LFC, and the sound /θ/ itself is also an intelligible pronunciation besides being a BrE and NAmE (nativelike) pronunciation.

The phonological and phonetic features that are opted for in the LFC, and which are also features of both BrE and NAmE, were not represented in the acceptability task since the acceptance of the pronunciations of such features could not indicate if a participant was more oriented to intelligibility or nativeness. An example of such features is the stressed aspirated /t/, pronounced as [tʰ], in a word like "tan". Also, where an intelligible feature has two nativeness equivalents in both NAmE and BrE, only one equivalent was kept in the acceptability task to maintain a balance

between the number of the intelligible and nativelike pronunciations based on the intelligibility and nativeness features; if two nativeness features were kept in the task, the number of the nativelike pronunciations would be more, and this may have affected the statistical analysis when comparing the responses to both the intelligible and nativelike pronunciations (see section 5.12.2 below for more details on the analysis of the acceptability task).

The 66 pronounced words in the acceptability task had the intelligibility and nativeness features represented in different positions of the words: i.e., word-initially, word-medially and word-finally where possible, since some features occur in one or two positions of the words only (e.g., the aspirated /p/ occurs word-initially or word-medially only). In addition, the intelligible alternatives that can replace some phonemes, according to the LFC, were all represented in the word examples in the acceptability task. For example, the dental fricatives /θ/ and /ð/ in English are commonly substituted in L2 by a restricted set of alternatives–i.e., /t/ and /d/, /s/ and /z/, or less commonly /f/ and /v/, and these substitutions are considered intelligible substitutions according to the LFC (see section 3.5, Chapter 3, for more details); thus these all were represented in the list of the word examples in the acceptability task (e.g., /[t̪]ɪk/ for "thick"; /enisɪŋ/ for "anything"; /maʊf/ for "mouth"; /[d̪]æn/ for "than"; /wɪvɪn/ for "within"; and /wɪz/ for "with").

Also as regards the pronounced words in the acceptability task, the quality of vowels and diphthongs must be used consistently regardless of which quality an L2 speaker uses according to the LFC (see section 3.5, Chapter 3); thus, the quality used in the acceptability task for the intelligible pronunciations of vowels and diphthongs was based on the variety of English used in the context of the study (Saudi Arabia). The selection of the quality was made based on the author's familiarity with the English spoken in that context. For example, the regular quality used in the context of the study for/aʊ/ in the word "house" is /ɑʊ/ (see Appendix M for the wordlist used for the acceptability task).

The words, whose pronunciations were recorded for the acceptability task (see above), were presented in the DMDX display system in random order. In the view of the author of this study, presenting the recorded words, whose pronunciations were either intelligible, nativelike or unacceptable, in random order could enhance the validity of the task. If the recorded words were presented in order, the participants may have been able to guess the goal of the task or may have understood the arrangement of the items in the task and thus responded in the same way to the intelligible, nativelike and unacceptable pronunciations. Also, the transcripts of the pronounced

words in the acceptability task were presented in the DMDX display system simultaneously with their recorded pronunciations. This procedure was undertaken to be sure that the participants would assess the pronunciation of the word whose transcript would be presented on the DMDX display system, rather than another similar one.

The script used to run both the visual and auditory stimuli on the DMDX display system was written following both the script examples and the help on writing the script available on the display system's website (http://www.u.arizona.edu/~kforster/dmdx/dmdx.htm). The DMDX display system presented the visual and auditory stimuli in pages.

The first part of the script included a line with symbols that were used to control the layout of the whole display, which included aspects such as the duration for the display of each page, the continuous or automatic running of the display rather than the need to press a key on the keyboard to continue, the input device (a keyboard, in case of this study), and the text size. Following this, the instructions on doing the acceptability task on the display system were given in the script in order to be presented on the first page of the DMDX display. In these instructions the participants were asked to judge the acceptability of the spoken words they would hear and see their transcripts on the following display pages. They were also instructed on the same page that they could decide the acceptability of the pronunciations by pressing the Right Shift key on their computer' keyboard to accept the pronunciations or the Left Shift key to reject them. They were also advised on the same page that they had only five seconds to respond to each spoken word. This limit of time was decided to control the duration of the task for all the participants. Furthermore, they were informed that they were required to do the task in a single session, without breaks (i.e., they could not pause the automatic task display during doing the task), and that they had about five minutes to complete the task. This information was given to the participants following the feedback received after piloting this task (see section 5.9.2 below for details on piloting the acceptability task). At the end of the first display page, the participants were instructed to press Spacebar key on their keyboard to begin the task since the first display page was not automatically running like the following pages.

Following the first display page, there were 66 display pages; each display page included the transcript of a word appearing at the centre of the page along with the pronunciation of the word. These display pages turn automatically after a participant sees the word and hears its pronunciation; they are timed to turn after five seconds following seeing and hearing the word. To run these pages on the DMDX display system, each page had one line script that included things like

the transcript of the word and its display order along with its pronunciation. The line script also indicated whether the words spoken were acceptable or not; this was indicated using the symbol "+" and "-", where the first symbol was used with a word whose pronunciation was intelligible and the latter was for nativelike and unacceptable pronunciations. If a participant accepted a spoken word with an intelligible pronunciation, the DMDX display system would consider this response correct, while if a participant accepted a spoken word with a nativelike or an unacceptable pronunciation, the DMDX display system would consider it wrong. The DMDX display system used error rate for indicating whether a response was correct or wrong: if the response was correct, the error rate would decrease, while if the response was wrong, the error rate would increase (see section 5.12.2 below for more details on the analysis of the acceptability task data).

Following the pages that displayed the words along with their pronunciations was the closing display page. The script used to run this page on the DMDX display system included a notice informing the participants that this was the end of task. On the same page, the participants were instructed to (a) end the task by pressing the Escape key on their keyboard, (b) copy the responses that appear in a text format following the ending of the task, (c) paste the responses copied into a Word document, (d) save them and finally (e) email them back to the researcher (see Appendix N for the script used for running the acceptability task on the DMDX display system).

5.4.4 Online Questionnaire

The third instrument to measure the EFL teachers' attitudes as regards intelligibility and nativeness in pronunciation assessment in this study was an online agree/disagree questionnaire. This instrument was used to answer the second sub-question of the main research question (RQ1b): do the EFL teachers demonstrate attitudes towards intelligibility or nativeness in an online questionnaire? Are the two groups of teachers different in this regard? The participants of the study were given 14 statements oriented to intelligibility or nativeness and asked to state whether they agree or disagree with them. It was hypothesised in this PhD study that the participants would be oriented to the type of statements they agreed with more: if they were oriented to intelligibility, they would agree more with the statements oriented to intelligibility, and vice versa.

The statements in the online questionnaire were constructed following Dörnyei's (2007) guidance for writing questionnaire items: e.g., he recommended that statements should be in simple language and expressing only one thought. All the statements of the questionnaire addressed

aspects of English pronunciation teaching and assessment. Seven of the statements were oriented to nativeness: e.g., "My students' pronunciations should sound nativelike (e.g., like a British/American English speaker)". Other seven statements were orientated to intelligibility: e.g., "It is unnecessary for my students to sound nativelike in order to be intelligible (understandable)". These two types of statements were ordered randomly through the questionnaire so that the participants would not be able to guess the purpose of the questionnaire. The whole questionnaire was made using Online Surveys website where it was available for the participants to access using a link.

In the online questionnaire the participants stated whether they agreed or disagreed with the statements using a seven-point scale, where degree 1 = strongly agree and degree 7 = strongly disagree. The seven-point scale was used since using as few items as five would make the respondents unable to express their real responses as the options are limited (Preston & Colman, 2000), while using as many items as nine would make the respondents unable to discriminate between such too many options (Miller, 1994). Given that the lowest degree (i.e., 1) indicated strong agreement and the highest degree (i.e., 7) indicated strong disagreement, the lower means of the participants' responses to the statements in the online questionnaire indicated stronger agreement with the statements, while the higher means indicated stronger disagreement (see Appendices O and P for a copy of the online questionnaire and intelligibility and nativeness statements; see section 5.11.2 below for details on the reliability of the online questionnaire).

5.4.5 Follow-up Interview

Follow-up interviews with some of the participants were conducted following data collection and analysis to further investigate the behaviour of these participants in the study tasks. These follow-up interviews were mainly aimed to answer the second parts of the third and fourth sub-questions of the main research question of this study (RQ1c and RQ1d): the reason behind any conflict between the participants' attitudes in the practical tasks and the online questionnaire, and the factors affecting the participants' attitudes (e.g., institutional assessment policy).

This PhD study was conducted to investigate the attitudes of EFL teachers to intelligibility and nativeness in pronunciation assessment. Two practical tasks in addition to an online questionnaire were used to investigate the attitudes (as detailed in the sections above). The attitudes emerged in the online questionnaire were expected to be reflected in the practical tasks. However, some participants did not meet the expectation. Thus, the follow-up interview was mainly

conducted to find out the reason behind this conflict between attitudes in the online questionnaire and behaviour in the practical tasks. In addition to this main purpose of the follow-up interview, the interview was used to find out any other factors affecting the participants' attitudes, besides age and experience which were investigated separately (see section 5.11.9 below).

5.4.5.1 The Participants of Follow-up Interview

Six participants were chosen for the follow-up interviews based on their behaviour in the study tasks: i.e., their behaviour in the practical tasks was different to their behaviour in the online questionnaire. Thus, the follow-up interview aimed to ascertain the reason behind this conflict between attitudes in the online questionnaire and behaviour in the practical tasks (as mentioned above). Firstly, the behaviour of each participant of the study was observed individually in order to determine their individual attitudes in every task. Firstly, the attitudes of each individual participant were decided after counting the responses of each individual participant as regards intelligibility and nativeness in each task following the analyses approaches and schemes explained in section 5.12 below. Next, following the hypotheses of the study tasks mentioned above (see sections 5.4.1.1, 5.4.3.1 and 5.4.4), the totals of the participants' responses in each task were used to decide the orientation of each individual participant. Following these procedures, the attitudes of all the study participants for all the tasks were determined. Next, the participants whose attitudes in the practical tasks were different to those in the online questionnaire were drawn for the follow-up interview. As a result, six participants were selected for the follow-up interview; these participants were all oriented to nativeness in the practical tasks, but they were oriented to intelligibility in online questionnaire.

5.4.5.2 The Structure of Follow-up Interview

The follow-up interview in this study was semi-structured; this type of interviews is conducted using prepared guiding questions but allows an interviewee to elaborate on certain issues of particular interest (Dörnyei, 2007). The prepared guiding questions in the follow-up interview in this study involved asking the participants of the interview about the reason behind any conflict between their behaviour in the practical tasks (i.e., the assessment of students' speech and acceptability tasks) and their attitudes in the online questionnaire. The participants' attitudes to intelligibility and nativeness in the online questionnaire should have been reflected in the participants' behaviour in the practical tasks: if the participants were oriented to nativeness in the

online questionnaire, they should have been oriented to nativeness in the practical tasks as well, and if they were oriented to intelligibility in the online questionnaire, they should have been oriented to intelligibility in the practical tasks. If this was not the case, the participants were asked about the reason behind this conflict in their attitudes in the practical tasks and the online questionnaire. This prepared guiding question helped answer the second part of RQ1c: the reason behind any conflict between the participants' attitudes in the practical tasks and the online questionnaire.

The prepared guiding questions also involved asking the participants of the interview about the factors that had affected their responses in the study tasks. This would help determine if their attitudes were affected by, for example, other speaking assessments such as the speaking assessment at KAU since the participants of the interview were working at KAU at the time of the study. The responses to this question were hoped to answer the second part of RQ1d: the factors affecting the participants' attitudes (e.g., institutional assessment policy). The attempt to determine the factors affecting the participants' attitudes would in return help teaching and assessment materials developers with controlling these factors. Given the relevance of intelligibility, particularly intelligibility based on the LFC, demonstrated in Chapter 2 and Chapter 3, the teaching and assessment materials developers can revise these factors in light of intelligibility so that they change EFL teachers' attitudes towards intelligibility.

In addition to preparing guiding questions, semi-structured interviews needed "extensive planning, from recruiting an appropriate sample […] to developing an interview guide" (Rolland et al., 2020, p. 280). Selecting an appropriate sample for the interview has been discussed previously above. As for the interview guide, it should consider the topics of interest to the researcher, and the questions in this guide should be open-ended questions in order to allow the interviewees formulate their responses "according to what is important to them" (Rolland et al., ibid., p. 280). In this vein, Kvale (2007) stated that the aim in interviews is to elicit responses in the interviewees' own words and to avoid leading questions when developing the interview guide. He further mentioned that the interview questions should not include assumptions or technical expressions (ibid.). As regards the overall structure of the interview guide, the interview should begin with an introductory question or "ice breaker" to make the interviewee feel relaxed and thus ready for the interview (ibid., 2020, p. 280). The following questions should be organised according to the topics in the interview, where each topic questions should be raised together (ibid.). At the end of the interview, there should be a closing question or an opportunity for the interviewee to talk about

anything of relevant to the interview main subject (ibid.). Accordingly, all these recommendations were followed in constructing the follow-up interviews of this study (see Table 5.3 below for the follow-up interview guide, including the interview guiding questions and the purpose of each guiding question).

In addition to the guiding questions, there were also specific follow-up questions that arose from the responses of the follow-up interview participants; they were mainly clarifying guiding questions. For example, a specific question that arose as a follow-up question to the first guiding question which was asking about the participants' opinion on the study tasks is the following: "Do you think they were appropriate for measuring teachers' attitudes to intelligibility and nativeness?" Another specific question arose as a clarifying question to the third guiding question asking about the factors affecting the participants' behaviour in the study tasks is the following: "I mean any background factors you think have affected your responses in the tasks such as particular experience, previous speaking assessments?" Also, another specific question came to clarify the second guiding question, which was preceded by a brief explanation for clarification is the following: "All the tasks were used to find out if you were oriented to intelligibility or nativeness. There is a framework called "the LFC" that was used to measure if you were oriented to either constructs. Based on your responses to the tasks, it seems you were more oriented to intelligibility in the online questionnaire but oriented to nativeness in the two practical tasks. So .. can you tell of any reasons behind such conflict?" Appendix Q shows the list of the follow-up questions.

Table 5.3: The follow-up interview guide, including the interview guiding questions and the purpose that each question serves

Guiding Question	Purpose of Guiding Question
1. You have done three tasks for my study: an assessment of students' speech task, an acceptability task, and an online questionnaire. What do you think of these tasks?	Serves as an ice-breaker
2. Your behaviour in the three tasks you have done show that you were oriented to nativeness/intelligibility in the online questionnaire but oriented to intelligibility/nativeness in the two practical tasks: the assessment of students' speech and acceptability tasks. Why?	Answers the second part of RQ1c: the reason behind any conflict between the participants' attitudes in the practical tasks and the online questionnaire.
3. What were the factors that affected your responses to the tasks? (A possible follow-up question: was your response affected by other speaking assessments such as the speaking assessment at KAU?	Answers the second part of RQ1d: the factors affecting the participants' attitudes.
4. Do you have anything else you want to add or say?	Serves as a closing question

5.5 Participants

A convenience sampling is used in this PhD study. According to Mackey et. al. (2005), convenience sampling is:

> [T]he selection of individuals who happen to be available for study. For instance, a researcher who wanted to compare the performance of two classes after using different review materials might select the two classes that require the review materials based on the curriculum. (p. 122)

Accordingly, the participants of this study came from Saudi universities, mainly KAU where the author of this PhD study works. Stratified random sampling was also used in this PhD study. According to Mackey et. al. (2005, p. 120), in stratified random sampling, "the proportions of the subgroups in the population are first determined, and then participants are randomly selected from within each stratum according to the established proportions". Accordingly, the population of EFL teachers with experience of teaching at a Saudi Arabian university were divided into two strata:

non-native and native English speakers (NNSs and NSs, respectively). Then the participants for this study were drawn from the two strata.

As mentioned above in section 1.1 (Chapter 1), classifying the speakers of English as a NNS and NS is a thorny issue. According to Davies (2003) and Cook (2007), the definition of a NS and NNS is a difficult issue and diverse aspects need to be considered when defining them such as psycholinguistic, linguistic and sociolinguistic perspectives. Thus, when grouping the teachers in this study into NNS and NS groups, only the countries where the teachers come from were considered. The EFL teachers who come from the countries of Inner Circle (see section 3.2.1, Chapter 3, for more details on this circle and the other circles) such as England, USA and Canada represented the stratum of NSs, and the EFL teachers who come from the Outer Circle such as India and the Expanding Circle such as Saudi Arabia represented the stratum of NNSs (see Table 5.4 below for the countries from where the participants come).

The participants of the PhD study were EFL teachers drawn mainly from KAU. Since the number of participants drawn from this university was not enough, and since further employment of participants was disrupted by the lockdown following the COVID-19 pandemic, a few more participants from other universities (other than KAU) were contacted by personal emails with the help of some friends of the author to reach these participants. Also, for the same reason, five female participants from KAU were included in the sample, although the plan in the first was to involve male EFL teachers only in order to control variables and because the physical access to female participants in Saudi Arabia was restricted because of religious and cultural boundaries. As a result, the participants were male and female EFL teachers who had experience of teaching EFL at a Saudi university, but the majority (N = 53 out of 57) came from KAU since these participants were the colleagues of the author, and it was convenient to approach them and access other relevant materials they used such as the speaking assessment policy. Some of these participants were, as far as the author was familiar with the practices on campus, responsible for preparing the rating scale for speaking assessment, and most of them were in charge of conducting the speaking assessment. The participation of these participants and all study participants was voluntary.

Out of the total of 57 participants of the study, 33 are NNS male EFL teachers and additional 19 NS male EFL teachers and five more NS female EFL teachers. The data from further two participants were removed from the study as they failed to complete all the tasks. The participants' ages ranged from 29 to 67 years old. Their first languages are mainly Arabic or English,

but few participants have other L1s besides Arabic and English. The NNS participants' L2 was mainly English with some participants having French as an additional L2, while the NS participants' L2s are Arabic, Urdu and Japanese. Their qualifications for teaching English ranged from bachelor's degree to PhD. Their estimated teaching experience of English at a university level ranged from one year to 40 years. The estimated times they had been speaking examiners at a university level ranged from less than 10 years to over 20. None of the participants reported specific qualifications for speaking or pronunciation assessment, and they did not report their familiarity with the LFC.

The participants were divided into two main groups based on the L1 variable: 33 NNS EFL teachers and 24 NS EFL teachers. Some of the other variables were investigated as regards their impact on the teachers' attitudes since such variables are reported in the literature to have a potential impact on language attitudes: age and teaching experience (as mentioned above in section 5.3). However, the other variables, e.g., gender, additional L1s and L2s, may have had an effect on the teachers' attitudes (although it is not reported in the literature), but they were not investigated regarding their effect due to the lack of a large enough sample of participants sharing the same background details. Background details about the participants are listed in table 5.4 below.

Table 5.4: Background details about the participants (not all information has been reported from the side of participants, so only information reported is listed in the table)

Background information	Number of participants
Age	Less than 40 = 26 (45.61%) 40-50 = 26 (45.61%) More than 50 = 5 (8.77%)
Gender	Male = 52 (91.22%) Female = 5 (8.77%)
Home countries (Nationalities)	22 = England (English) 14 = Tunisia (Tunisian) 13 = Saudi Arabia (Saudi) 3 = Egypt (Egyptian) 2 = Jordan (Jordanian) 1 = Pakistan (Pakistani) 1 = Canada (Canadian) 1 = USA (American)
Native languages	Arabic = 32 (56.14%) English = 21 (36.84%) Arabic & Urdu = 1 (1.75%) English & Urdu = 1 (1.75%) English & Hindi = 1 (1.75%) English & Korean = 1 (1.75%)
Second languages	English = 19 (33.33%) English & French = 14 (24.56%) Urdu = 4 (7.01%) Arabic = 2 (3.5%) Japanese = 1 (1.75%) Arabic & Urdu = 1 (1.75%)
Qualification for teaching English	MA TESOL/Linguistics = 45 (78.94%) PhD TESOL/Linguistics = 9 (15.78%) BA TESOL/Linguistics = 3 (5.26%)
Number of years teaching at university level	Less than 10 = 31 (54.38%) 10-20 = 16 (28.07%) More than 20 = 7 (12.28%)
Number of times assessing speaking at university level	> 10 = 23 (40.35%) 10-20 = 12 (21.05%) < 20 = 17 (29.82%)

5.6 Procedures for Collecting Practical Tasks and Online Questionnaire Data

The original plan for data collection was to meet the participants of this study on campus, at KAU, and then collect the data from them. Unfortunately, when the author of this study travelled to Saudi Arabia to collect the data from the participants, this coincided with the summer term at KAU, and thus there were only a few teachers available on campus. As a result, three teachers were asked to collect the data on the behalf of the author. These three teachers accepted to collect the data in first, but the contact with two of them could not be maintained, and the third teacher expressed his inability to collect the data on his own. Thus, some of the data were collected personally by the author from the participants through communication via email and phone. However, the number of the participants was not enough for statistical analysis. As a result, more data was planned to be collected through personal meetings with other participants at other campuses in Saudi Arabia; thus, the author of this study travelled again to Saudi Arabia. Unfortunately, the teaching at universities in Saudi Arabia was turned online due to the COVID-19 pandemic this time, so more data were again collected through communication via email and phone with more participants.

The participants of this study were contacted via their university email addresses or their WhatsApp numbers and asked to participate in the research voluntarily. Following their agreement to participate in the study, they were given a brief explanation of the research tasks, and they were sent the files of the research tasks including the ethical approval forms. They were instructed to do the assessment of students' speech task first, then the acceptability task, and finally the online questionnaire. This order of the tasks was decided as it was assumed that the assessment of students' speech task was the least task that would indicate the purpose of the study, followed by the acceptability task and then the online questionnaire. If the participants did the online questionnaire first, for example, it was highly possible they would guess the purpose of the study and thus this would affect their behaviour when they were doing the practical tasks.

The assessment of students' speech task along the personal background questionnaire were sent to the participants in one Word file along with the six audio recordings in a zipped file. The recordings were numbered from one to six according to their order in the assessment of students' speech task. Prior to data collection, no instruction was given to the participants regarding the phonetic and phonological terminology. However, examples and brief explanations next to each phonetic or phonological term were provided to make the terminology clear to the participants (see

Appendix K). Also, the terminology used in the task was supposed to be of common knowledge to the participants of this study as their qualifications were a BA, MA or PhD in TESOL/Linguistics; a basic knowledge of phonetics and phonology terminology was therefore assumed.

The acceptability task was sent to the participants in a zipped file including the recordings of the spoken words (in MP3 format) and the scripts (in RFT format) to run the task. Along with this zipped file, the participants were also sent a short video on how to download the DMDX display system and how to run the acceptability task on it. Finally, the participants were sent a link to the online questionnaire, and they completed it online.

5.7 Procedures for Conducting Follow-up Interviews

As mentioned above (see section 5.4.5.2 above), the follow-up interviews in this study were semi-structured interviews. The planning of semi-structured interviews should include aspects such as deciding on how many sessions required for the interview and deciding on the medium used for conducting the interview (e.g., face-to-face or virtual; Rolland et al., 2020). The planning should also involve deciding on the location where the interview is conducted and how long the interview takes (ibid.). It should further involve deciding on the way for recording the interview (ibid.). In this regard, Rolland et al. (ibid.) recommended recording the interview using voice recorders since other techniques to keep a record of the conversation such as note-taking is too difficult to provide enough detail. Accordingly, the follow-up interviews in this study were planned and then conducted according to the planning details mentioned above. They were conducted in one session (according to Rolland et al. (ibid.), interviews are usually held in a single session). Also, the interviews in this study were conducted virtually given that the interviews took place during the COVID-19 pandemic. In this regard, Rolland et al. (ibid.) mentioned that interviews are commonly conducted face-to-face; however, technology still provides an opportunity for virtual interviews conducted in both spoken and written form (O'Connor et al., 2008). Also, the follow-up interviews in this study took from 10-15 minutes, and they were recorded using two digital voice recorders.

Finally, the follow-up interviews in this study were conducted following the procedures recommended by the literature for conducting an interview. For example, according to Rolland et al. (2020), interviewees need to be informed that the interview is recorded. They also need to be informed about the purpose of the study in advance of interviewing them (ibid.). Accordingly, in

advance of the follow-up interviews in this study, permission was obtained from the participants of the interviews for recording the interviews, and the purpose of the interviews was explained to them. In addition to these procedures, the follow-up interviews were conducted with the participants one at a time, and they were conducted in English.

5.8 Transcription of Follow-up Interviews

The interviews were transcribed following conducting and recording them. The transcription conventions used for the interviews of this study were adopted from Schiffrin (1994, pp. 422-433; see Appendices R and S for a sample of the conventions used and the transcriptions of the follow-up interviews). The content of the interviews of this study was more important than paralanguage, and thus a broader form of transcription was used.

5.9 Piloting Research Instruments

A pilot study was conducted in advance of the data collection for the main study, where the study tasks were all piloted. However, the data collected during the pilot study was not included in the data of the main study. Ethical approval was gained from the University of Reading, where the author of this study is a PhD researcher, and from the participants of the pilot study in advance of the data collection for the pilot study. The ethical approval forms that were used for the main study had also been used for the pilot study (see section 5.10 below for more details).

5.9.1 Piloting Assessment of Students' Speech Task

The assessment of students' speech task was piloted with two participants. These two participants were colleagues of the author of this study; they were PhD students of Applied Linguistics at the University of Reading, and all of them are English NNSs and they were EFL teachers in their home countries. They were sent the assessment of students' speech task along with the recordings of the L2 learners via email, and they were asked to do the task. The assessment of students' speech task was sent in a Word document where the participants could insert their responses in the file and then save it and send it back.

Piloting the assessment of students' speech task has benefited in shortening the task. The task during the pilot study included the assessment of the recordings of two reading passages besides one free speech. After piloting, the task was shortened to include the recordings of one

reading passage only besides one free speech. This change to the task was undertaken following the feedback of one of the participants of the pilot study; this participant said that the task was too long and laborious. She said the task had taken her almost an hour and 20 minutes to finish. After removing one recording out of three recordings which are almost equally long, it was estimated that the assessment task of the two remaining recordings would take almost 50 minutes to finish. Also, since the participants were not asked to do the task in one session, it was hoped that the task would not be too long and cognitively demanding for the participants.

5.9.2 Piloting Acceptability Task

The acceptability task was piloted with two participants who were met individually in the UK, and they completed the acceptability task following the guidance of the author of this study. One of the two participants was a colleague of the author; he was a PhD student of Applied Linguistics at the University of Reading. The other participant was attending a pre-sessional English language course; he was attending this course that was taught at the University of Reading in order to prepare for his study of an MA in Applied Linguistics. The author of this study had the acceptability task ready on a laptop when he met the participants, and he took the two participants through the task.

Piloting the acceptability task has helped in providing more information to the instructions of the task in order to make the participants prepared better for the task. For example, following the recommendation of one of the participants who did the acceptability task, the duration of the task, which was about 5 minutes, and the need to complete the task in one session without breaks during the task were all indicated in the task instructions.

5.10 Ethical Consideration

Ethical issues concerning data collection from and about individuals must inevitably be carefully considered (Punch, 2014). Accordingly, in the current study the policies and related administrative processes as regards the research ethics were pursued. Firstly, an application form to conduct this research was submitted to the University of Reading's research ethics committee. After the approval was obtained from the university's ethics committee, approval for collecting the data from KAU was sought from vice presidency for graduate studies and scientific research at KAU.

Approval was obtained to collect the data from the EFL teachers and students at the university under the supervision of some staff at the university (see Appendix D).

Cohen et al. (2018) discussed the issue of "deception" in research; they stated that the purpose and context of the research need to be made clear to the participants and they should not be "deceived" or misinformed about the procedures of the research (ibid.). Thus, the students, whose speech was recorded for the assessment of students' speech task (see section 5.4.1.2 below for more details), were informed by the teacher who collected the recordings on behalf of the author (see section 5.4.1.3 below) and through a consent form about the main purpose of the research. They were informed that the purpose is to explore EFL teachers' attitudes in pronunciation assessment, but without giving details in order not to affect the data collection procedures. These students were informed that they would be asked to read two passages in English, which would be used as audio stimuli in the research project (see section 5.4.1.3 below). They were reassured that their recordings would be accessible only to the participants in the study, the teacher who collected the data, the author and his supervisors, and then they would be destroyed as soon as the research has been concluded. They were further informed that their recordings would be presented anonymously to the participants and that their privacy and confidentiality would be carefully observed, and all data would be stored on a password-protected computer or in a locked drawer. Further, the students were informed that they could withdraw from participating in the project at any stage. They were also informed that in case they had any further queries or they needed any further clarifications on anything in the research project, they could contact the supervisor of the author at the email address they were provided in the form sheet. Finally, they were asked to sign the consent form and information sheet that contained explanation of the purpose and procedures of this PhD research (see Appendices E and F).

Similarly, the teachers who participated in this PhD study were informed by the author of this PhD study through a consent form about the main purpose of the study. They were also informed that their responses would be used or accessed only by the author, his supervisors and the research assessment committee. Also, they were informed that their names would be collected for the study to assist analysis, but no names or personal details would be revealed in the research or subsequent publications. Like the students, the teachers were reassured that their privacy and confidentiality would be carefully observed, and all the data collected would be stored on a password-protected computer or in a locked drawer. They were also informed that they had the option to withdraw from the study at any stage of data collection. They were also advised that if

they had any queries about this PhD study or they needed any more information, they could contact the supervisor of the author at the address they were provided. They were finally asked to sign the consent form and an information sheet containing explanation of the purpose and procedures of this PhD study (see Appendices G and H).

5.11 Reliability of Study Instruments

In the subsections below the inter-reliability of the assessment of students' speech task along with the reliability of the online questionnaire are discussed.

5.11.1 Inter-Rater Reliability of Assessment of Students' Speech Task

After the author of this study (the first rater) has completed counting the responses of each individual participant as regards intelligibility and nativeness in the assessment of students' speech task following the analyses approaches and schemes explained in section 5.12.1 below, a second rater was recruited to carry out an inter-rater reliability measurement. She was a PhD student of Applied Linguistics (a colleague of the author of this study at the University of Reading). She was firstly emailed and given the participants' responses in attached Word files. The assessment of students' speech task was explained to the second rater in detail, including the design of the task mentioned above (see section 5.4.1.1). The second rater was further asked to count the responses of each individual participant as regards intelligibility and nativeness in the assessment of students' speech task. She was given the analyses approaches and schemes mentioned above and asked to follow them when counting the responses.

After receiving back the totals of the responses the second rater had had found for each participant on intelligibility and nativeness, these totals along with the totals found by the author of this study (the first rater) were inserted into SPSS for inter-rater reliability analysis. This analysis was made to measure the degree of agreement between the two raters. Cronbach's Alpha analysis, which is the most common method to test internal consistency reliability (Stemler, 2004) was used for this purpose. The result of this analysis shows high Cronbach's Alpha coefficients for the totals found by the two raters for both intelligibility and nativeness (totals of intelligibility responses = .865, totals of nativeness responses = .859). Thus, this suggests a high degree of agreement between the two raters, and thus this indicates a high reliability level of the analysis made by the author of this study.

5.11.2 Reliability of Online Questionnaire

The responses of the online questionnaire were used to analyse the reliability of the questionnaire. Cronbach's alpha analysis was used to measure the reliability of the online questionnaire. The result of this analysis showed an acceptable degree of reliability (.717).

5.12 Data Analysis

5.12.1 Analysis and Scoring of Assessment of Students' Speech Task

The analysis scheme of the data of the assessment of students' speech task was developed by the author of this study without referring to the literature given that there are no previous studies that used a task similar to the assessment of students' speech task in this study (as far as the author of this study is aware).

In the assessment of students' speech task, the participants of this study (the NNS and NS EFL teachers) were asked to assess L2 learners' speech based on phonological and phonetic features they were provided. The phonological and phonetic features totalled eighteen, where ten of them were intelligibility features (i.e., the core features of the LFC and which are important for intelligibility in ELF speech, see section 3.5, Chapter 3) and the remaining eight were nativeness features (i.e., the non-core features of the LFC and which are not important for intelligibility in ELF speech, see section 3.5, Chapter 3). The participants checked these features against six recordings from the three L2 learners; the features were used for each recording. If a participant found any feature problematic in any of the recordings, s/he gave examples for the features s/he found problematic from the recordings. The participants were also asked in a separate section to provide any other phonological or phonetic features they found problematic in the recordings (section 5.4.1.1, Chapter 5, for more details on this).

Firstly, in the analysis of the participants' responses to the assessment of students' speech task, the examples the participants gave to the phonological, and the phonetic features listed in the task were verified; they were checked to find out if each example included a phonological or phonetic problem matched the feature to which the example was given for. This was done to ensure that the participants had understood the phonological and phonetic features listed in the assessment of students' speech task, and thus they gave examples that included phonological or phonetic problems matching the features. It was also done to ensure the participants were taking

the task seriously (see section 5.4.1.1, Chapter 5, for more details). Next, both the intelligibility and nativeness features the participants detected in the recordings were counted separately. The calculation of the features involved the phonological and phonetic features listed in the assessment of students' speech task and the ones given in the second section of the task (where the participants were asked to give any further phonological or phonetic features they found problematic; see section 4.6.1.1, Chapter 5). The phonological or phonetic features the participants gave in the second section of the task were considered only if they were part of the core and non-core features of the LFC. The core features are the ones important for intelligibility in ELF speech and thus they are intelligibility features, and the non-core features are not important for intelligibility in ELF speech and thus they are nativeness features (see section 3.5, Chapter 3, for more details on the LFC). The maximum score that could be obtained on the intelligibility items in the assessment of students' speech task was ten since the intelligibility features listed in the task (and which are the core features of the LFC) totalled ten features. The maximum score that could be obtained on the nativeness items in the assessment of students' speech task was eight since the nativeness features listed in the task (and which are the non-core features of the LFC) totalled eight features.

In the calculation of the phonological and phonetic features listed in the assessment of students' speech task, the following two points were considered:

a. Since the participants used the phonological and phonetic features (which include both intelligibility and nativeness features) to assess six recordings (as mentioned above), if a feature was detected once in one recording, this feature was counted only for one time when calculating the features for each participant. Counting the same feature for more than one time if detected by a participant for more than one time would provide misleading findings as regards the participant's attitudes. According to the hypothesis of the assessment of students' speech task, if a participant was more oriented to intelligibility than nativeness, s/he would be more interested in the intelligibility features than the nativeness features when assessing, and thus s/he would detect more intelligibility features in the recordings than the nativeness features, and vice versa (see section 5.4.1.1, Chapter 5). Thus, if a participant detected the same nativeness feature in the recordings for three times, for example, while s/he detected three different intelligibility features in the recordings, it would be misleading to conclude that the participant in this case was equally oriented to intelligibility and nativeness since s/he detected the same number of features for both intelligibility and nativeness. In this case, it was more accurate to conclude that

the participant was more oriented to intelligibility than nativeness; the participant was more interested in intelligibility than nativeness as s/he detected three different intelligibility features, while s/he detected only one nativeness feature.

b. All the examples given by the participants and were incorrectly spelled were considered correct examples of the phonological and phonetic features detected because these examples were still identifiable. For example, the word “cal” was provided as an example of the mispronunciation of the dark /l/, and thus it would be easy to guess that the word misspelled here was “call”, particularly if this word was in the recording assessed and there was no other word more similar.

A quantitative approach to analysis was adopted for analysing the data of the assessment of students’ speech task. The data of this task included the phonological and phonetic features identified by the participants as problematic. The phonological and phonetic features detected were firstly classified as either intelligibility or nativeness features according to the LFC. It should be noted that the intelligibility features are the core features of the LFC and the nativeness features are the non-core features of the LFC. The core features of the LFC are the features important for intelligibility in ELF speech and thus they are intelligibility features. The non-core features are not important for intelligibility in ELF speech and thus they are nativeness features (see section 3.5, Chapter 3). Next, a quantitative approach was conducted for calculating the intelligibility and nativeness features detected by the participants separately following the analysis scheme mentioned above. The intelligibility features detected by each participant as problematic were calculated and each participant was given a score out of ten (i.e., the maximum score on the intelligibility items, as mentioned above) according to the total of the intelligibility features each participant detected as problematic. Also, the nativeness features detected by each participant as problematic were calculated and each participant was given a score out of eight (i.e., the maximum score on the nativeness items, as mentioned above) according to the total of the nativeness features each participant detected as problematic. Since the intelligibility and nativeness features detected as problematic were calculated using two different scales, where the intelligibility features detected were calculated on a scale out of ten while the nativeness features detected were calculated on a scale out of eight, the scores calculated for each participant on the assessment of students’ speech task were turned into percentages before running the analysis. For example, if a participant detected three intelligibility features out of ten and four nativeness features out of eight, these scores were turned into 30% and 50%, respectively. Next, the percentages of the scores of the

intelligibility and nativeness features detected were inserted into SPSS and compared using statistical tests.

5.12.2 Analysis and Scoring of Acceptability Task

In the acceptability task, the participants of the study were asked to judge the acceptability of the pronunciations of particular words. These pronunciations were three types: (1) intelligible pronunciations, which were built on the intelligibility features; (2) nativelike pronunciations, which were built on the nativeness features; and (3) totally wrong pronunciations, which were used as a distractor in order that the participants could not guess the purpose of the task and to check if the participants were taking the task seriously.

The task was presented using a computer software called DMDX. Using the software, the participants were shown the words on screen while hearing their pronunciations, and then they gave their responses using a keyboard. At the end of the task, the responses appear on the computer's screen in a text format. The participants were asked to copy the responses text and email it to the author of this study (see section 5.4.3 above for more details on the acceptability task).

The responses texts that the participants copied following the end of the acceptability task included the error rates for the words pronounced. As mentioned in section 5.4.3 above, in the script used to run the acceptability task on DMDX software, the intelligible, nativelike and totally wrong pronunciations were indicated using the symbols "+" and "-" to know when a participant accepted a pronunciation or rejected it. The symbol "+" was used for the words whose pronunciations are intelligible, and if a participant accepted the pronunciations of these words, the error rate would decrease, and vice versa: if the participant rejected them, the error rate would increase. However, the symbol "-" was used for the words whose pronunciations are nativelike or wrong, and if a participant accepted the pronunciations of these words, the error rate would increase, and vice versa: if the participant rejected them, the error rate would decrease. This does not mean that if a participant accepted a nativelike pronunciation, it is an erroneous response, but this would only help differentiate between accepted and rejected pronunciations. Accordingly, (a) the accepted intelligible pronunciations (the ones decreasing the error rate) were counted separately, (b) the accepted nativelike pronunciations (the ones increasing the error rate) were counted separately, and (c) the totally wrong pronunciations (the ones increasing the error rate)

were counted separately for each participant (using the error rate to differentiate between them). Thus, each participant had a score out of 22 for each category as there were 22 intelligible pronunciations, 22 nativelike pronunciations, and 22 totally wrong pronunciations (as mentioned above in section 5.4.3). However, the unaccepted pronunciations were not counted since they would give the same conclusions as the accepted pronunciations. For example, if a participant rejected fewer intelligible pronunciations, s/he would accept more pronunciations of this type, and vice versa. Also, if s/he rejected fewer nativelike pronunciations, s/he would accept more pronunciations of this type, and vice versa. The scores for the totally wrong pronunciations, which were used only as a distractor and to ensure the participants were taking the task seriously (as mentioned above), suggest that all the participants took the task seriously since almost all of them were rated unacceptable by the participants. The scores for the intelligible pronunciations accepted which were calculated separately for each participant, and the scores for the nativelike pronunciations accepted, which were also calculated separately for each participant were then inserted into SPSS and compared using statistical tests.

5.12.3 Analysis and Scoring of Online Questionnaire

A quantitative approach was conducted for the data of the online questionnaire. The data of the online questionnaire included the degrees on a 7-point scale used in the questionnaire (for commentary on why a 7-point scale was adopted, see section 5.4.4). The participants of the study were asked in the online questionnaire to state the degrees of their agreement/disagreement with statements using a scale of 7 points, where 1 = strongly agree and 7 = strongly disagree. The statements in the online questionnaire were two types: intelligibility statements, which were oriented to intelligibility, and nativeness statements, which were oriented to nativeness (see section 5.4.4 above for more details on the online questionnaire). The means of the participants' responses to the intelligibility and nativeness statements were calculated, where each participant had a score out of 7 for his/her responses to the intelligibility statements and a score out of 7 for his/her responses to the nativeness statements as the scale for the responses was 7 points. Next, the scores were inserted into SPSS and compared using statistical tests.

5.12.4 Comparison within Groups

According to Pallant (2010), "[a] paired-samples t-test is used when you have only one group of people (or companies, or machines etc.) and you collect data from them on two different

occasions or under two different conditions" (p. 243). However, there are some assumptions that the data need to meet in order to be valid for analysis using a paired-samples t-test (Pallant, ibid.). Pallant (ibid.) described the assumptions as follows:

- Level of measurement: the scale of measurement for the variables should be interval or ratio (continuous, e.g., scores);
- Random sampling;
- Independence of participants: participants must be independent of one another. Measurements for one participant do not affect measurements for any other participant.
- Normal distribution: this means that the data near the average are more frequent in occurrence than the data that are away from the average. However, according to Tabachnick and Fidell (2013, p. 251), a sample size of at least 20 should ensure robustness, and thus there is no need for normality test.

The data obtained from all the study tasks (i.e., the assessment of students' speech task, the acceptability task and the online questionnaire) met the above-mentioned assumptions; thus, several paired-samples t-tests using SPSS were conducted in order to compare the responses of both the NNS and NS EFL teachers to both intelligibility and nativeness items in the study tasks. In detail, paired-samples t-tests were conducted in order to compare the following: (a) the scores of the NNS EFL teachers on the intelligibility and nativeness items in the assessment of students' speech task; (b) the scores of the NS EFL teachers on the intelligibility and nativeness items in the assessment of students' speech task; (c) the totals of the intelligible and nativelike pronunciations accepted by the NNS EFL teachers in the acceptability task; (d) the totals of the intelligible and nativelike pronunciations accepted by the NS EFL teachers in the acceptability task; (e) the NNS EFL teachers' means of their responses to the intelligibility and nativeness statements in the online questionnaire; and (f) the NS EFL teachers' means of their responses to the intelligibility and nativeness statements in the online questionnaire. These paired-samples t-tests helped answer the first parts of RQ1a, RQ1b and RQ1c: are the NNS and NS EFL teachers oriented to intelligibility or nativeness when doing the study tasks?

In the output of the paired-samples t-test, Sig. (2-tailed), which is the probability (p) value, shows the significance of difference between scores (Pallant, 2010). If this value is less than .05 (e.g., .04, .01, .001), it can be concluded that there is a significant difference between scores (ibid.). If the

value is above .05 (e.g., .06, .10), there is no significant difference between the two scores (ibid.). Having found that there is a significant difference, there needs to determine which set of scores is higher and which set is lower to interpret the scores accordingly (ibid.). To determine this, the output of the paired-samples t-test gives the mean scores for each of the two sets of scores, and thus they can be compared to find out which set is higher and which is lower.

Following a paired-samples t-test, effect sizes need to be calculated. According to Pallant (2010), effect size (also known as "strength of association") can assess the importance of the findings of a paired-sample t-test. According to Tabachnick and Fidell (2013, p. 54), an effect size is a set of statistics that show the relative magnitude of the differences between means. There are a number of different effect size statistics; one of the most commonly used statistics is *Cohen's d*, which "presents difference between groups in terms of standard deviation units" (Pallant, ibid., p. 210). According to Cohen (1988), the guidelines for interpreting the value of *Cohen's d* are as follows: .20 = small effect, .50 = moderate effect, and .80 = large effect. However, according to Plonsky and Oswald (2014), L2 researchers should adopt the new field-specific benchmarks of small (d = .40), medium (d = .70), and large (d = 1.00). Thus, in this PhD study, Plonsky and Oswald's (2014) benchmarks were used for interpreting the values of effect sizes. Finally, the effect sizes were calculated using the effect size calculator available at https://www.socscistatistics.com/effectsize/default3.aspx.

5.12.5 Comparison across Groups

Several independent-samples t-tests were conducted using SPSS to compare both the NNS and NS EFL teachers in terms of their responses to both intelligibility and nativeness items in the study tasks and thus answer the second parts of RQ1a, RQ1b and RQ1c: Are the NNS and NS EFL teachers different in terms of their attitudes to intelligibility and nativeness when an acceptability task and an online questionnaire?

According to Pallant (2010), an independent-samples t-test is used to compare the mean score, on some continuous variable, for two different groups of participants. In this study, there were two groups (i.e., the NNS and NS EFL teachers) for the data of the study tasks, and each group had two scores on two different continuous variables (i.e., intelligibility and nativeness). Thus, the two groups were compared in terms of their responses to intelligibility items separately and to nativeness items separately in the study tasks. In detail, independent-samples t-tests were

conducted in order to compare the following: (a) the totals of the intelligibility items detected by the two groups of teachers as problematic; (b) the totals of the nativeness items detected by the two groups of teachers as problematic; (c) the totals of the intelligible pronunciations accepted by the two groups of teachers in the acceptability task; (d) the totals of the nativelike pronunciations accepted by the two groups in the acceptability task; (e) the two groups' means of their responses to the intelligibility statements in the online questionnaire; and (f) the two groups' means of their responses to the nativeness statements in the online questionnaire.

However, the data need to meet five assumptions in order that the independent-samples t-test produces a valid result. Four of these assumptions are the same required for the paired-samples t-test mentioned above (see section 5.12.4 above): level of measurement, random sampling, independence of observations, normal distribution. These four assumptions were met by the data of the study tasks, as mentioned above (see section 5.12.4 above). In addition to these four, the independent-samples t-test requires the assumption of homogeneity of variance. According to Pallant (2010), homogeneity of variance means that the samples of a study are drawn from populations of equal variances; in other words, the variability of scores for each of the study groups is similar (ibid.). To test this assumption, SPSS performs Levene's test for equality of variances as part of the independent-samples t-test, where the results of the Levene's test are presented in the output of the independent-samples t-test (ibid.). If the significance value of the Levene's test is less than .05, this suggests that the variances for the two groups are not equal, and thus this violates the assumption of homogeneity of variance (ibid.). However, the independent-samples t-test is reasonably robust to violations of the assumption of homogeneity of variance, provided the size of the study groups is reasonably similar (like the case of this study: the study includes a group of 33 NNS EFL teachers and 24 NS EFL teachers; ibid.). The output of the independent-samples t-test presents two sets of results: one set is for the situations where the assumption of homogeneity of variance is not violated, and the other set is for the situations where it is violated (ibid.). Thus, the output of the independent-samples t-test is interpreted based on whichever set of results is appropriate for the study data.

Given the details mentioned in the preceding paragraph, independent-samples t-tests were used for the data of the study tasks in the way mentioned above. Following the independent-samples t-tests, effect sizes were calculated for the output of these t-tests. These analyses were conducted and interpreted in the same way as the ones conducted for the output of the paired-samples t-tests (see section 5.12.1 above).

5.12.6 Correlation Analyses for Attitudes in Practical Tasks and Online Questionnaire

Correlation analyses were conducted in this study in order to answer the first part of RQ1d: are the NNS and NS EFL teachers' attitudes in the practical tasks (i.e., the assessment of students' speech task and the acceptability task) and the online questionnaire in harmony or in conflict? The correlation analyses were conducted to investigate whether there were relationships between the dependent variables (attitudes to intelligibility and nativeness) in the practical tasks and the online questionnaire: i.e., if increases in the dependent variables (attitudes to intelligibility and nativeness) in the practical tasks were associated with increases in the online questionnaire, and vice versa. The strength and direction of the correlations, whether a positive or a negative correlation, would help answer the first part of RQ1c mentioned above.

The correlation analysis indicates the size of the value of correlation coefficient (Pallant, 2010). This value indicates the strength of the relationship between the variables analysed (ibid.). The value can range from -1.00 to 1.00 (ibid.). A negative sign in front of the correlation coefficient value indicates there is a negative correlation between the variables: i.e., high scores on one variable are correlated with low scores on the other variable (ibid.). A value of 0 indicates that the correlation analysis shows no relationship between the variables at all; a value of 1.00 indicates there is a perfect positive correlation between the variables; and a value of -1.00 indicates a perfect negative correlation between the variables (ibid.). The strength of values is interpreted as follows: small r = .10 to .29, medium r = .30 to .49, and large r = .50 to 1.0 (Cohen, 1988). The significance of correlation, which suggests the reliability of correlation findings, is indicated by the p-value, and interpreted as mentioned previously (see section 5.12.4 above).

5.12.7 Analyses of Effects of Participants' Backgrounds on Attitudes

Correlation analyses were conducted in this study in order to answer the first part of RQ1d: are the NNS and NS EFL teachers' attitudes in the study tasks affected by their personal background information (e.g., age and experience)? Correlation analyses were conducted to investigate the relationships between the independent variables of this study (i.e., age, teaching experience and speaking assessment experience) and the dependent variables (i.e., attitudes to intelligibility and nativeness). In other words, the correlation analyses were used to investigate if the independent variables and the dependent variables correlated so that if the independent

variables increased, the dependent variables would increase, or vice versa (interpretation of correlation analyses is explained above in section 5.12.6).

The effect of age, teaching experience and speaking assessment experience on the attitudes was measured in this study since age and teaching experience have been recognised by some of the literature as factors affecting language attitudes (see section 4.2.1, Chapter 4). Speaking assessment experience has not been identified as a factor affecting language attitudes but, since this study is looking at language attitudes in pronunciation assessment and since teaching experience usually involves assessment experience, speaking assessment experience was investigated. Thus, it was decided to investigate whether this was the case in this study.

5.12.8 Analysis of Follow-up Interviews

Deductive analysis for the data of the follow-up interviews was conducted. In this process, the second parts of RQ1d and RQ1e were determined as categories: the reason behind any conflict between the participants' attitudes in the practical tasks and the online questionnaire, and the factors affecting the participants' attitudes (e.g., institutional assessment policy). Next, any connections in the data of the follow-up interviews were mapped to the categories.

5.13 Summary of Chapter

This chapter has described the methodology used for this exploratory study. It has discussed the research design for this study. It also described the participants of this study and the instruments for collecting the data. This was followed by a description of the procedures used for collection of the data and then the process of piloting the study instruments. Finally, the analysis of the data was discussed.

CHAPTER 6: RESULTS

6.1 Introduction

This chapter begins with presenting the results of the study tasks: the practical tasks (i.e., the assessment of students' speech task and the acceptability task) and the online questionnaire. Next, the chapter presents the findings of the correlation analyses conducted for the results of the practical tasks and the online questionnaire. This is followed by the findings of the correlation analyses conducted for the independent variables and the results of the study tasks. Finally, the chapter concludes with the results of the follow-up interviews.

6.2 Comparison within and across Groups

Below are the results of the t-tests used to compare between and within groups (i.e., the NNS and NS EFL teachers) in the study tasks. It should firstly be noted that in the assessment of students' speech task the core features of the LFC are the features important for intelligibility in ELF speech and thus they are classed as intelligibility features in the analysis. Also, the non-core features of the LFC in the same task are not important for intelligibility in ELF speech and thus they are classed as nativeness features (see section 5.4.1.1, Chapter 5). In the assessment of students' speech task, it was hypothesised in this PhD study that if a participant was more oriented to intelligibility, s/he would tend to focus more on the intelligibility features in the assessment of students' speech task and thus s/he would identify more intelligibility features as problematic than nativeness features, and vice versa; that is to say, if a participant was more oriented to nativeness, s/he would tend to focus more on the nativeness features and thus s/he would identify more nativeness features as problematic than intelligibility items (see section 5.4.1.1, Chapter 5). It should be noted that the maximum totals of the intelligibility and nativeness items detected as problematic in the assessment of students' speech task (which are ten and eight, respectively) were converted into percentages, and thus the maximum totals are 100 for the intelligibility and nativeness items detected as problematic (see section 5.12.1, Chapter 5).

The acceptability task included a list of 66 words with their scripts representing 22 intelligible pronunciations, 22 nativelike pronunciations, and 22 totally wrong pronunciations (see section 5.4.3, Chapter 5). The totally wrong pronunciations were developed based on avoiding the common substitutions of sounds allowed by the LFC and based on variations between the wrong

pronunciations and the correct ones in terms of place and manner of articulation (see section 5.4.3, Chapter 5). These pronunciations were used only as a distractor and to ensure the participants were taking the task seriously (see section 5.12.2, Chapter 5). The findings regarding these pronunciations suggest that all the participants took the task seriously since almost all of them were rated unacceptable by the participants. Thus, no more analysis of these pronunciations was provided in this chapter. The intelligible pronunciations in the acceptability task were based on the core features of the LFC, and the nativelike pronunciations were based on the non-core features of the LFC. In the acceptability task, it was hypothesised in this PhD study that a participant would accept more intelligible or nativelike pronunciations according to their attitudes: if a participant was more oriented to intelligibility, s/he would accept more intelligible pronunciations than nativelike ones, and vice versa: i.e., if s/he was more oriented to nativeness, s/he would accept more nativelike pronunciations than intelligible ones (see section 5.4.3, Chapter 5). It should be noted that the maximum totals of the intelligible and nativelike pronunciations accepted in the acceptability task are 22 per each (see section 5.12.2, Chapter 5).

In the online questionnaire, the maximum totals of the degrees of agreement/disagreement with the intelligibility and nativeness statements are 7 per each (see section 5.12.3, Chapter 5). It should be noted below that Sig. (2-tailed), which is the probability (p) value, shows the significance of difference between scores (Pallant, 2010). If this value is less than .05 (e.g., .04, .01, .001), this indicates a significant difference between scores, while if it is above .05 (e.g., .06, .10), this indicates no significant difference between scores (ibid., see section 5.12.4, Chapter 5).

6.2.1 Comparison within Groups in Assessment of Students' Speech Task

As mentioned in section 5.12.4 (Chapter 5), two paired-samples t-tests were conducted and effect sizes were calculated to help answer the first part of RQ1a of the main research question (see section 5.2, Chapter 5, for the main research question): Are the NNS and NS EFL teachers more oriented to intelligibility or nativeness when doing an assessment of students' speech task? In this case, a statistically significant difference within the groups in the assessment of students' speech task would suggest the teachers have detected one category of pronunciation features (either intelligibility or nativeness) as problematic more than the other. Table 6.1 below provides a summary of the outputs of the paired-samples t-tests and effect sizes for the data of the assessment of students' speech task.

Table 6.1: A summary of the outputs of the paired-samples t-tests conducted for the responses of the two groups to both intelligibility and nativeness items in the assessment of students' speech task

Group	Pair	Mean of responses	SD	Mean difference	95% Confidence interval of the Difference		t	Df	Sig. (2-tailed)	Cohen's d
					Lower	Upper				
NNS	Intelligibility	43.03	22.56	-7.72	-22.38	6.93	1.07	32	.29	.28
	Nativeness	50.75	31.24							
NS	Intelligibility	29.58	20.10	-2.70	-14.16	8.75	0.48	23	.62	.12
	Nativeness	32.29	24.70							

*p<0.05, **p<0.01, ***p<0.001

The output of the paired-samples t-test conducted for the responses of the NNS EFL teachers showed there was not a statistically significant difference between the totals of the intelligibility items detected (M = 43.03, SD = 22.56) and the nativeness items detected (M = 50.75, SD = 31.24), t (32) = 1.07, p = .29 (two-tailed). The mean difference was -7.72 with a 95% confidence interval ranging from -22.38 to 6.93. Cohen's d (.28) showed a negligible effect size. Similarly, the output of the paired-samples t-test conducted for the responses of the NS EFL teachers failed to show a statistically significant difference between the totals of the intelligibility items detected (M = 29.58, SD = 20.10) and the nativeness items detected (M = 32.29, SD = 24.70), t (23) = .48, p = .62 (two-tailed). The mean difference was -2.70 with a 95% confidence interval ranging from -14.16 to 8.75. The Cohen's d (.12) indicated a negligible effect size. These results suggest there were no differences between the number of intelligibility and nativeness features detected as problematic by both groups of teachers in the assessment of students' speech task. According to the hypothesis of the assessment of students' speech task mentioned above, these results imply that the attitudes of both groups of teachers failed to show more orientation to intelligibility or nativeness.

6.2.2 Comparison within Groups in Acceptability Task

As mentioned in section 5.12.4 (Chapter 5), two paired-samples t-tests were conducted and effect sizes were calculated to help answer the first part of RQ1b of the main research question: Are the study groups more oriented to intelligibility or nativeness when doing an acceptability task? Again, a statistically significant difference within the groups in the acceptability task would suggest the total of the intelligible or nativelike pronunciations accepted by a group is more than the other

total. Table 6.2 below provides a summary of the outputs of the paired-samples t-tests and effect sizes for the data of the acceptability task.

Table 6.2: A summary of the outputs of the paired-samples t-tests conducted for the responses of the two groups to both intelligibility and nativeness items in the acceptability task

Group	Pair	Mean of responses	SD	Mean difference	95% Confidence interval of the Difference		t	Df	Sig. (2-tailed)	Cohen's d
					Lower	Upper				
NNS	Intelligibility	18.33	3.04	.60	-.86	2.07	.83	32	.408	.20
	Nativeness	17.73	2.91							
NS	Intelligibility	17.38	2.99	-.66	-2.13	.80	-.94	23	.357	.20
	Nativeness	18.04	3.40							

*p<0.05, **p<0.01, ***p<0.001

The output of the paired-samples t-test conducted for the responses of the NNS EFL teachers in the acceptability task showed there was not a statistically significant difference between the totals of the intelligible pronunciations accepted (M = 18.33, SD = 3.04) and the nativelike pronunciations accepted (M = 17.73, SD = 2.91), *t* (32) = .83, *p* = .408 (two-tailed). The mean difference was .60 with a 95% confidence interval ranging from -.86 to 2.07. Cohen's d (.20) showed a negligible effect size. Similarly, the output of the paired-samples t-test conducted for the responses of the NNS EFL teachers failed to show a statistically significant difference between the totals of the intelligible pronunciations accepted (M = 17.38, SD = 2.99) and the nativelike pronunciations accepted (M = 18.04, SD = 3.40), *t* (23) = -.94, *p* = .357 (two-tailed). The mean difference was -.66 with a 95% confidence interval ranging from -2.13 to .80. The Cohen's d (.20) indicated a negligible effect size. These results suggest that there were no differences between the number of the intelligible and nativelike pronunciations accepted by both groups of teachers in the acceptability task. According to the hypothesis of the acceptability task mentioned above (see section 6.2), these results indicate that the attitudes of both groups of teachers in the acceptability task failed to show more oriented to either intelligibility or nativeness.

6.2.3 Comparison within Groups in Online Questionnaire

As mentioned in section 5.12.4 (Chapter 5), two paired-samples t-tests were conducted and effect sizes were calculated to help answer the first part of RQ1c of the main research question (see section 5.2, Chapter 5, for the main research question): Are the study groups more oriented to

intelligibility or nativeness when doing an online questionnaire? Achieving a statistically significant difference within the groups in the online questionnaire would suggest the degrees of agreement/disagreement with one category of the statements (the intelligibility or nativeness statements) are higher than the degrees of agreement/disagreement with the other category. Table 6.3 below provides a summary of the outputs of the paired-samples t-tests and effect sizes for the data of the online questionnaire.

Table 6.3: A summary of the outputs of the paired-samples t-tests conducted for the responses of the two groups to both intelligibility and nativeness statements in the online questionnaire

Group	Pair	Mean of responses	SD	Mean difference	95% Confidence interval of the Difference		t	Df	Sig. (2-tailed)	Cohen's d
					Lower	Upper				
NNS	Intelligibility	3.54	.77	-.15	-.58	.27	-.71	32	.479	.18
	Nativeness	3.69	.85							
NS	Intelligibility	2.87	.92	-.75	-1.25	.24	-3.07	23	.005**	.81
	Nativeness	3.62	.90							

*p<0.05, **p<0.01, ***p<0.001

The output of the paired-samples t-test conducted for the responses of the NNS EFL teachers showed there was not a statistically significant difference between the means of their responses to the intelligibility statements (M = 3.54, SD = .77) and the nativeness statements (M = 3.69, SD = .85), *t* (32) = -.71, *p* = .479 (two-tailed). The mean difference was -.15 with a 95% confidence interval ranging from -.58 to .27. Cohen's d (.18) showed that the difference was negligible. This result suggests that there was no difference between the NNS EFL teachers' degrees of agreement/disagreement with both the intelligibility and nativeness statements in the online questionnaire. This in return suggests that the attitudes of this group of teachers failed to show more orientation to intelligibility or nativeness. However, the output of the paired-samples t-test conducted for the responses of the NS EFL teachers showed there was a statistically significant difference between the means of their responses to the intelligibility statements (M = 2.87, SD = .92) and the nativeness statements (M = 3.62, SD = .90), *t* (23) = -3.07, *p* = .005 (two-tailed). The mean difference was -.75 with a 95% confidence interval ranging from -1.25 to .24. Cohen's d (.81) indicated a medium effect size. Given that the lower means of the participants' responses to the statements in the online questionnaire indicate stronger agreement with the statements, and vice versa (see section 5.4.4, Chapter 5), the NS group agreed with the intelligibility statements more

than the nativeness statements. This suggests that the attitudes of the NS EFL teachers were oriented more to intelligibility than nativeness.

6.3 Comparison across Groups

6.3.1 Comparison across Groups in Assessment of Students' Speech Task

As mentioned in section 5.12.5, Chapter 5, two independent-samples t-tests were conducted and effect sizes were calculated to help answer the second part of RQ1a: Are the study groups (i.e., the NNS and NS EFL teachers) different from each other in terms of their attitudes towards both intelligibility and nativeness when doing the assessment of students' speech task? Achieving a statistically significant difference between the groups in the assessment of students' speech task would suggest that the number of intelligibility and/or nativeness items that one group detected as problematic is more than the number of intelligibility and/or nativeness items detected by the other group as problematic. Table 6.4 below presents a summary of the outputs of the independent-samples t-tests and effect sizes for the data of the assessment of students' speech task.

Table 6.4: A summary of the outputs of the independent-samples t-tests conducted for the responses of the two groups to both the intelligibility and nativeness items in the assessment of students' speech task

Group	Pair	Mean of responses	SD	Mean difference	95% Confidence interval of the Difference		t	Df	Sig. (2-tailed)	Cohen's d
					Lower	Upper				
NNS	Intelligibility	43.03	22.57	10.53	-1.19	22.25	1.8	55	.07	.48
NS		32.50	20.69							
NNS	Nativeness	43.62	22.39	10.33	-1.50	22.15	1.19	55	.07	.49
NS		32.29	24.70							

*p<0.05, **p<0.01, ***p<0.001

The output of the independent-samples t-test conducted for the responses of the two groups regarding the intelligibility items showed no statistically significant difference between the NNS group (M = 43.03, SD = 22.57) and the NS group (M = 32.50, SD = 20.69), t (55) = 1.8, p = .07 (two-tailed). The mean difference was 10.53 with a 95% confidence interval ranging from -1.19 to 22.25. Cohen's d (.48) showed that the effect size of the difference was negligible. Similarly, the output of the independent-samples t-test conducted for the responses of the two groups regarding the nativeness items detected in the assessment of students' speech task revealed that there was not

a statistically significant difference between the NNS group (M = 43.62, SD = 22.39) and the NS group (M = 32.29, SD = 24.70), *t* (55) = 1.19, *p* = .07 (two-tailed). The mean difference was 10.33 with a 95% confidence interval ranging from -1.50 to 22.15. Cohen's d (.49) indicated that the difference was negligible. These results suggest that there were no differences between the two groups of teachers in terms of the number of the intelligibility and nativeness features they detected as problematic in the assessment of students' speech task. According to the hypothesis of the assessment of students' speech task (see section 6.2 above), these results imply that both groups of teachers were equal in terms of their attitudes towards intelligibility and nativeness.

6.3.2 Comparison across Groups in Acceptability Task

As mentioned in section 5.12.5, Chapter 5, two independent-samples t-tests were conducted and effect sizes were calculated to help answer the second part of RQ1b: Are the study groups (i.e., the NNS and NS EFL teachers) different from each other in terms of their attitudes towards both intelligibility and nativeness when doing the acceptability task? A statistically significant difference between the groups in the acceptability task would suggest that the number of the intelligible and/or nativelike pronunciations accepted by one group is more than the number of the intelligible and/or nativelike pronunciations accepted by the other group. Table 6.5 below presents a summary of the outputs of the independent-samples t-tests and effect sizes for the data of the acceptability task.

Table 6.5: A summary of the outputs of the independent-samples t-tests conducted for the responses of the two groups to both the intelligibility and nativeness items in the acceptability task

Group	Pair	Mean of responses	SD	Mean difference	95% Confidence interval of the Difference		t	Df	Sig. (2-tailed)	Cohen's d
					Lower	Upper				
NNS	Intelligibility	18.33	3.04	.95	-.66	2.58	1.18	55	.243	.31
NS		17.38	2.99							
NNS	Nativeness	17.73	2.91	-.31	-1.99	1.36	-.37	55	.710	.09
NS		18.04	3.40							

*p<0.05, **p<0.01, ***p<0.001

The output of the independent-samples t-test conducted for the responses of the two groups regarding the intelligible pronunciations showed no statistically significant difference between the NNS group (M = 18.33, SD = 3.04) and the NS group (M = 17.38, SD = 2.99), t (55) =

1.18, *p* = .243 (two-tailed). The mean difference was .95 with a 95% confidence interval ranging from -.66 to 2.58. Cohen's d (.31) showed that the effect size of the difference was negligible. Similarly, the output of the independent-samples t-test conducted for the responses of the two groups regarding the nativelike pronunciations revealed that there was not a statistically significant difference between the NNS group (M = 17.73, SD = 2.91) and the NS group (M = 18.04, SD = 3.40), *t* (55) = -.37, *p* = .710 (two-tailed). The mean difference was -.31 with a 95% confidence interval ranging from -1.99 to 1.36. Cohen's d (.09) indicated that the difference was negligible. These results suggest that there were no differences between both groups of teachers in terms of the number of the intelligible and nativelike pronunciations they accepted in the acceptability task. According to the hypothesis of the acceptability task (see section 6.2 above), these results imply that the attitudes of both groups of teachers were the same in terms of their orientation towards intelligibility and nativeness.

6.3.3 Comparison across Groups in Online Questionnaire

As mentioned in section 5.12.5, Chapter 5, two independent-samples t-tests and effect sizes were calculated to help answer the second part of RQ1c: Are the study groups (i.e., the NNS and NS EFL teachers) different from each other in terms of their attitudes towards both intelligibility and nativeness when doing the online questionnaire? Achieving a statistically significant difference between the groups in the online questionnaire suggests that the degrees of agreement/disagreement with the intelligibility and/or nativeness statements of one group are higher than the degrees of the other group. Table 6.6 below provides a summary of the outputs of the independent-samples t-tests and effect sizes for the data of the online questionnaire.

Table 6.6: A summary of the outputs of the independent-samples t-tests conducted for the responses of the two groups to both the intelligibility and nativeness items in the online questionnaire

Group	Pair	Mean of responses	SD	Mean difference	95% Confidence interval of the Difference		t	Df	Sig. (2-tailed)	Cohen's d
					Lower	Upper				
NNS	Intelligibility	3.54	.77	.63	.17	1.09	2.76	54	.004**	.73
NS		2.91	.92							
NNS	Nativeness	3.69	.85	-.07	-.39	.54	.30	55	.760	.07
NS		3.62	.90							

*p<0.05, **p<0.01, ***p<0.001

The output of the independent-samples t-test conducted for the responses of the two groups regarding the intelligibility statements showed a statistically significant difference between the NNS group (M = 3.54, SD = .77) and the NS group (M = 2.91, SD = .92), t (54) = 2.76, p = .004 (two-tailed). The mean difference was .63 with a 95% confidence interval ranging from .17 to 1.09. Cohen's d (.73) showed that the effect size of the difference was medium. Given that the lower means of the participants' responses to the statements in the online questionnaire indicate stronger agreement with the statements, and vice versa (see section 5.4.4, Chapter 5), this result suggests that the NS group agreed with the intelligibility statements more than the NNS group. In return, this result suggests that the attitudes of the NS group were more oriented to intelligibility than the NNS group. However, the output of the independent-samples t-test conducted for the responses of the two groups regarding the nativeness statements revealed that there was not a statistically significant difference between the NNS group (M = 3.69, SD = .852) and the NS group (M = 3.62, SD = .90), t (55) = .30, p = .760 (two-tailed). The mean difference was -.07 with a 95% confidence interval ranging from -.39 to .54. Cohen's d (.07) indicated that the difference was negligible. This result suggests that there was no difference between both groups' degrees of agreement/disagreement with the nativeness statements in the online questionnaire. This in return suggests that both groups of teachers were equal in terms of their attitudes towards intelligibility and nativeness in the online questionnaire.

6.4 Correlations between Teachers' Attitudes in Practical Tasks and Online Questionnaire

As mentioned in section 5.12.6 (Chapter 5), correlation analyses were conducted to answer the first part of RQ1d: Are the study groups' attitudes to intelligibility and nativeness in the practical tasks (i.e., the assessment of students' speech task and the acceptability task) and the online questionnaire in harmony or in conflict? Thus, Spearman Correlation analyses were used to investigate whether there were relationships between the dependent variables (the attitudes to intelligibility and nativeness) in the practical tasks and the online questionnaire.

While correlations cannot be taken to suggest causation or one variable affecting the other, the positive correlations would suggest that an increase in one variable was associated with an increase in the other variable, and vice versa. Negative correlations, however, suggest that an increase in one variable was associated with a decrease in the other variable, and vice versa. Insignificant correlations suggest that an increase or a decrease in one variable was not associated with an increase or a decrease in the other variable.

It should be noted that all the results of the correlation analyses reported in the subsections below were not statistically significant, and where the variables were correlated, the correlations were weak.

6.4.1 Correlations between NNS EFL Teachers' Attitudes towards Intelligibility and Nativeness

Firstly, to determine the relationships between the totals of the intelligibility and nativeness features detected by the NNS EFL teachers in the assessment of students' speech task and the means of their responses to the intelligibility and nativeness statements in the online questionnaire, Spearman Correlation analyses were run using SPSS. The output of the analyses showed that the correlation between the totals of the intelligibility features detected and the means of their responses to the intelligibility statements was not statistically significant (r = -.095; p = .600). Also, insignificant and a weak correlation was observed between the totals of the nativeness features detected and the means of their responses to the nativeness statements (r = -.132; p = .464).

Two more Spearman Correlation analyses were run to determine the relationships between the totals of the intelligible and nativelike pronunciations accepted by the NNS EFL teachers in the acceptability task and the means of their responses to the intelligibility and nativeness statements in the online questionnaire. The output of the analyses showed a weak and insignificant correlation between the totals of the intelligible pronunciations accepted and the means of their responses to the intelligibility statements (r = .224; p = .211). The correlation between the totals of the nativelike pronunciations accepted and the means of their responses to the nativeness statements was not statistically significant (r = -.005; p = .976; see Table 6.7 below for a summary of the outputs of the correlation analyses for the dependent variables of the NNS EFL teachers).

All the results mentioned above suggest that the intelligibility and nativeness features detected by the NNS EFL teachers as problematic in the assessment of students' speech task and the intelligible and nativelike pronunciations they accepted in the acceptability task were not necessarily associated with their responses to the intelligibility and nativeness statements in the online questionnaire. These results in return indicate the attitudes of the NNS EFL teachers in the

online questionnaire were not necessarily reflected in their behaviour in both the assessment of students' speech task and the acceptability task.

Table 6.7: A summary of the outputs of the correlation analyses for the dependent variables of the NNS EFL teachers

		Intelligibility responses of online questionnaire	Nativeness responses of online questionnaire
Intelligibility responses of assessment of students' speech task	Correlation Coefficient	-.095	---
	Sig. (2-tailed)	.600	---
Nativeness responses of assessment of students' speech task	Correlation Coefficient	---	-.132
	Sig. (2-tailed)	---	.464
Intelligibility responses of acceptability task	Correlation Coefficient	.224	---
	Sig. (2-tailed)	.211	---
Nativeness responses of acceptability task	Correlation Coefficient	---	-.005
	Sig. (2-tailed)	---	.976

6.4.2 Correlations between NS EFL Teachers' Attitudes towards Intelligibility and Nativeness

Firstly, to determine the relationships between the totals of the intelligibility and nativeness features detected by the NS EFL teachers in the assessment of students' speech task and the means of their responses to the intelligibility and nativeness statements in the online questionnaire, Spearman's Correlation analyses were run using SPSS. The output of the analyses showed a weak and insignificant correlation between the totals of the intelligibility features detected and the means of their responses to the intelligibility statements (r = -.115; p = .592). Similarly, a weak and insignificant correlation was observed between the totals of the nativeness features detected and the means of their responses to the nativeness statements (r = -.167; p = .435).

Two more Spearman's Correlation analyses were run to determine the relationships between the totals of the intelligible and nativelike pronunciations accepted by the NS EFL teachers

in the acceptability task and the means of their responses to the intelligibility and nativeness statements in the online questionnaire. The output of the analyses showed that the correlation between the totals of the intelligible pronunciations accepted and the means of their responses to the intelligibility statements was not statistically significant (r = -.057; p = .791). Similarly, the correlation between the totals of the nativelike pronunciations accepted and the means of their responses to the nativeness statements was not statistically significant (r = .019; p = .930; see Table 6.8 below for a summary of the outputs of the correlation analyses for the dependent variables of the NS EFL teachers).

All the above-mentioned results imply that the intelligibility and nativeness items detected by the NS EFL teachers as problematic in the assessment of students' speech task and the intelligible and nativelike pronunciations they accepted in the acceptability task were not necessarily associated with their responses to the intelligibility and nativeness statements in the online questionnaire. In return, these results indicate that the attitudes of the NS EFL teachers in the online questionnaire were not necessarily reflected in their behaviour in both the assessment of students' speech task and the acceptability task.

Table 6.8: A summary of the outputs of the correlation analyses for the dependent variables of the NS EFL teachers

		Intelligibility responses of online questionnaire	Nativeness responses of online questionnaire
Intelligibility responses of assessment of students' speech task	Correlation Coefficient	-.115	---
	Sig. (2-tailed)	.592	---
Nativeness responses of assessment of students' speech task	Correlation Coefficient	---	-.167
	Sig. (2-tailed)	---	.435
Intelligibility responses of acceptability task	Correlation Coefficient	-.057	---
	Sig. (2-tailed)	.791	---
Intelligibility responses of assessment of students' speech task	Correlation Coefficient	---	.019
	Sig. (2-tailed)	---	.930

6.5 Effects of Participants' Backgrounds on Attitudes

As mentioned in section 5.12.7 (Chapter 5), Spearman's Correlation analyses were conducted to answer the first part of RQ1d: are the participants' attitudes in the study tasks affected by some of the personal background factors (e.g., age, teaching experience and speaking assessment experience; the independent variables)? Thus, the correlation analyses were used to investigate whether there are relationships between the dependent variables (the attitudes to intelligibility and nativeness) in the study tasks and the independent variables (e.g., age, teaching experience and speaking assessment experience). The dependent variables included (a) the scores the participants obtained on the assessment of student's speech task (out of 100% for both the intelligibility and nativeness items; see section 5.12.1, Chapter 5), (b) the scores obtained on the acceptability task (out of 22 for both the intelligible and nativelike pronunciations; see section 5.12.2, Chapter 5), and (c) the scores obtained on the online questionnaire (out of 7 for both the intelligibility and nativeness statements; see section 5.12.3, Chapter 5). The independent variables included the participants' ages (ranged from 29-67 years), their teaching experience at a university level (ranged from one year to 40 years), and the times they had been speaking examiners at a university level (ranged from less than 10 years to over 20; see section 5.5 and Table 5.4, Chapter 5). Their scores, ages, teaching experience (number of years) and their speaking assessment experience (number of times they had been speaking examiners at a university level) were inserted into SPSS in order to conduct correlation analyses between these values.

It should be noted that all the results of the correlation analyses reported in the subsections below were not statistically significant, and almost all the analyses showed non-correlations or weak correlations.

6.5.1 Effects of NNS EFL Teachers' Backgrounds

Firstly, to determine the relationships between the totals of the intelligibility and nativeness features detected by the NNS EFL teachers in the assessment of students' speech task and their ages, and teaching experience and speaking assessment experience, Spearman's Correlation analyses were run using SPSS. The outputs of these analyses showed weak and insignificant correlations between the intelligibility features detected and the teaching experience and speaking assessment experience (r = .144 and .171, respectively; p = 340 and .630, respectively). The analyses also showed the correlation between the intelligibility features detected

by the NNS EFL teachers and their ages was not statistically significant (r = .087; p = .423). The outputs of the analyses also showed weak and insignificant correlations between the nativeness features the NNS EFL teachers detected in the assessment of students' speech task and their ages, teaching experience and speaking assessment experience (r = -.173, -.137 and -.108, respectively; p = .335, .447 and .551, respectively).

Two further Spearman's Correlation analyses were conducted to determine the relationships between the totals of the intelligible and nativelike pronunciations accepted by the NNS EFL teachers in the acceptability task and their ages and teaching experience and speaking assessment experience. The outputs of these analyses showed the correlations between the intelligible pronunciations accepted and age, teaching experience and speaking assessment experience were not statistically significant (r = -.061, -.068 and -.017, respectively; p = .735, .706 and .926, respectively). The outputs of the analyses also showed a weak and insignificant correlation between the nativelike pronunciations accepted by the NNS teachers in the acceptability task and their speaking assessment experience (r = -.143; p = .428). The analyses further revealed the correlations between the nativelike pronunciations accepted by the NNS teachers and their ages and teaching experience were not statistically significant (r = -.022 and -.062, respectively; p = .901 and .733, respectively).

Finally, two Spearman's Correlation analyses were conducted to determine the relationships between the means of the NNS EFL teachers' responses to the intelligibility and nativeness statements in the online questionnaire and their ages and their teaching experience and speaking assessment experience. The outputs of these analyses showed a moderate but insignificant correlation between the means of the NNS EFL teachers' responses to the intelligibility statements and their ages (r = .325; p = .152). Weak and insignificant correlations between the means of the NNS EFL teachers' responses to the intelligibility statements and their teaching experience and speaking assessment experience were also observed (r = .283 and .255, respectively; p = .064 and .111, respectively). Similarly, the outputs of the analyses also showed a moderate but insignificant correlation between the means of the NNS EFL teachers' responses to the nativeness statements and their speaking assessment experience (r = .309; p = .080). There were also weak and insignificant correlations between the means of the NNS EFL teachers' responses to the nativeness statements and their ages and teaching experience (r = .154 and .208, respectively; p = .392 and .245; see Table 6.9 below for a summary of the outputs of correlation analyses for the dependent and independent variables of the NNS EFL teachers).

All the above-mentioned results suggest that age, teaching experience and speaking assessment experience were not necessarily associated with the NNS EFL teachers' behaviour in the study tasks. Thus, the NNS EFL teachers' attitudes in the study tasks were unlikely to have been affected by their ages, teaching experience or their speaking assessment experience.

Table 6.9: A summary of the outputs of correlation analyses for the dependent and independent variables of the NNS EFL teachers

		Age	Teaching experience	Speaking assessment experience
Intelligibility responses of assessment of students' speech task	Correlation Coefficient	.144	.171	.087
	Sig. (2-tailed)	423	.340	.630
Nativeness responses of assessment of students' speech task	Correlation Coefficient	-.173	-.137	-.108
	Sig. (2-tailed)	.335	.447	.551
Intelligibility responses of acceptability task	Correlation Coefficient	-.061	-.068	-.017
	Sig. (2-tailed)	.735	.706	.926
Nativeness responses of acceptability task	Correlation Coefficient	-.143	-.022	-.062
	Sig. (2-tailed)	.901	.733	.428
Intelligibility responses of online questionnaire	Correlation Coefficient	.325	.283	.255
	Sig. (2-tailed)	.064	.111	.152
Nativeness responses of online questionnaire	Correlation Coefficient	.154	.208	.309
	Sig. (2-tailed)	.392	.245	.080

6.5.2 Effects of NS EFL Teachers' Backgrounds

Firstly, to determine the relationships between the totals of the intelligibility and nativeness features detected by the NS EFL teachers in the assessment of students' speech task and their ages, teaching experience and speaking assessment experience, Spearman's Correlation analyses were run using SPSS. The outputs of these analyses showed a weak and insignificant correlation between the intelligibility features detected by the NS EFL teachers and their speaking

assessment experience (r = -.134; p = .533). The analyses also showed the correlations between the intelligibility features they detected and their ages and teaching experience were not statistically significant (r = .023 and -.061, respectively; p = .913 and .778, respectively). Furthermore, the outputs of the analyses showed a weak and insignificant correlation between the nativeness features the NS EFL teachers detected in the assessment of students' speech task and their speaking assessment experience (r = -.200; p = .349). Further, the correlations between the nativeness features they detected and their ages and teaching experience were not statistically significant (r = -.051 and -.038, respectively; p = .813 and .858, respectively).

Two further Spearman's Correlation analyses were conducted to determine the relationships between the totals of the intelligible and nativelike pronunciations accepted by the NS EFL teachers in the acceptability task and their ages and teaching experience and speaking assessment experience. The outputs of these analyses showed weak and insignificant correlations between the intelligible pronunciations accepted by the NS EFL teachers and their ages and teaching experience (r = .179 and .132, respectively; p = .404 and .538, respectively). Furthermore, the correlation between the intelligible pronunciations accepted by the NS EFL teachers and their speaking assessment experience was not statistically significant (r = .059; p = .784). Also, the outputs of the analyses showed a weak and insignificant correlation between the nativelike pronunciations accepted by the NS EFL teachers and their ages, teaching experience and speaking assessment experience (r = -.182, -.170 and -.124, respectively; p = .394, .426 and .565, respectively).

Finally, two Spearman's Correlation analyses were conducted to determine the relationships between the means of the NS EFL teachers' responses to the intelligibility and nativeness statements in the online questionnaire and their ages, teaching experience and speaking assessment experience. The outputs of these analyses showed weak and insignificant correlations between the means of the NS EFL teachers' responses to the intelligibility statements and their ages and teaching experience (r = .212 and .197, respectively; p = .320 and .357, respectively). Moreover, the outputs of the analyses showed the correlation between the means of the NS EFL teachers' responses to the intelligibility statements and their speaking assessment experience was not statistically significant (r = .096; p = .656). Similarly, the correlations between the means of the NS EFL teachers' responses to the nativeness statements and their ages, and their teaching experience and speaking assessment experience were not statistically significant (r = .064, .030 and -.002, respectively; p = .767, .889 and .992, respectively; see Table 6.10 below for a summary of the

outputs of correlation analyses for the dependent and independent variables of the NS EFL teachers).

Table 6.10: A summary of the outputs of correlation analyses for the dependent and independent variables of the NS EFL teachers

		Age	Teaching experience	Speaking assessment experience
Intelligibility responses of assessment of students' speech task	Correlation Coefficient	-.134	.023	-.061
	Sig. (2-tailed)	.913	.778	.533
Nativeness responses of assessment of students' speech task	Correlation Coefficient	-.200	-.051	-.038
	Sig. (2-tailed)	.813	.858	.349
Intelligibility responses of acceptability task	Correlation Coefficient	.179	.132	.059
	Sig. (2-tailed)	.404	.538	.784
Nativeness responses of acceptability task	Correlation Coefficient	-.182	-.170	-.124
	Sig. (2-tailed)	.394	.426	.565
Intelligibility responses of online questionnaire	Correlation Coefficient	.212	.197	.096
	Sig. (2-tailed)	.320	.357	.656
Nativeness responses of online questionnaire	Correlation Coefficient	.064	.030	-.002
	Sig. (2-tailed)	.767	.889	.992

All the above-mentioned results suggest that age, teaching experience and speaking assessment experience were not necessarily associated with the NS EFL teachers' behaviour in the study tasks. These results suggest that the NS EFL teachers' attitudes in the study tasks were unlikely to have been affected by their ages, teaching experience or their speaking assessment experience.

6.6 Results of Follow-up Interviews

As mentioned in section 5.4.5 (Chapter 5), interviews with six of the participants of the main study were conducted to further investigate their behaviour in the study tasks. Although the data gathered from their responses was not fulsome in quantity nor always particularly useful in terms of insight, I feel it is worth reporting some of them. Firstly, since these participants were oriented to intelligibility in the online questionnaire but oriented to nativeness in the practical tasks (i.e., the assessment of students' speech task and the acceptability task), they were asked about the reason behind having different attitudes across the study tasks. Two participants stated they were not sure why they had done so. However, they explained that their answers in the online questionnaire were based on their beliefs that intelligibility is the measure for speech performance, but in the practical tasks they thought they were looking for intelligible speech. For example, one of the two participants said:

> *I think intelligibility is the benchmark for speech performance. So .. I think I answered the questionnaire based on this .. in the other tasks .. I think I was looking for intelligible speech.*

Three other participants said they did not know why they behaved differently across the study tasks and were unable to give an explanation. For example, one of these participants said:

> *I really don't have an answer for this question. It may, however, be due to my lack of awareness of the terminology. I had to look up the terms "nativeness" and "intelligibility" principles back then.*

The remaining participant stated that it is the same to be oriented to either intelligibility or nativeness since the two terms are interrelated. He said:

> *I believe that .. nativeness and intelligibility are somehow interrelated. Being able to speak nativelike mandates being intelligible. So .. focusing on nativelike pronunciation when teaching students, for instance, .. will necessarily lead to intelligibility.*

Next, the participants were asked about the factors that affected their responses to the study tasks. One participant said he was not aware of any factors affecting his responses, including other speaking assessments such as the speaking assessment at KAU (the context of the study, and where this participant was working at the time of this study). Two other participants thought that

maybe their experience as L2 English learners at school was the reason behind their responses. For example, one participant said:

> *Maybe the education I had in school affected my opinions. When I was a student, especially at university, the teachers, the textbooks, the audios for listening .. all of these taught the British accent.*

One participant thought that his experience of learning English in general may have affected his attitudes. He said:

> *Maybe because my learning experience .. all the things I learned was in American English*

The remaining two participants thought that the speaking assessment policies at KAU (where they were working at the time of this study) may have affected their responses. One of these participants thought even the speaking assessment policies where he had worked previously may have also influenced his responses to the study tasks. This participant said when he was asked if his responses were affected by other speaking assessments such as the speaking assessment at KAU:

> *I think I was affected by these assessments .. , and also the assessments when I was teaching in my country.*

6.7 Summary of Chapter

This chapter has firstly presented the results of the study tasks. Next, the chapter has presented the findings of the correlation analyses conducted for the results of the practical tasks and the online questionnaire and the correlation analyses conducted for the independent variables and the results of the study tasks. Finally, the chapter has presented the results of the follow-up interviews.

CHAPTER 7: DISCUSSION

7.1 Introduction

This chapter discusses the findings in relation to the research questions and in light of previous literature. It begins with the findings on the attitudes of the EFL teachers towards intelligibility and nativeness in the study tasks. It firstly introduces a summary of the major findings of EFL teachers' attitudes towards nativeness and intelligibility in the study tasks. Then, it discusses the attitudes of the EFL teachers study tasks towards nativeness and intelligibility. Next, this chapter discusses the correlations between the teachers' attitudes in the practical tasks (i.e., the assessment of students' speech task and the acceptability task) and the online questionnaire. Then, the effect of the independent variables on the teachers' attitudes is discussed. The other factors that may have affected the EFL teachers' attitudes, and which were mentioned by the teachers in the follow-up interview, are then discussed. Other findings that have revealed following in-depth analysis of the data are also discussed. The findings of the follow-up interview are discussed in different places of the chapter and where relevant. Finally, the chapter concludes with the limitations of the study.

Although the attitudes of the EFL teachers in this PhD study and the correlations between their attitudes in the online questionnaire and the practical tasks (i.e., the assessment of students' speech task and the acceptability task) may be similar to those in previous research reviewed in section 4.2.1 (Chapter 4; also as mentioned in the sections below), the attitudes in the current PhD study still make important contributions to the literature since they were found in conditions different to most of such research. This PhD study (a) has looked at the attitudes of the EFL teachers in pronunciation assessment; (b) it has investigated the attitudes through direct and indirect methods (i.e., an online questionnaire and practical tasks; see section 5.3, Chapter 5); (c) it considered the NNS and NS EFL teachers separately when investigating the attitudes; (d) adopted the LFC as a framework for intelligibility; and (e) it considered the attitudes in a new context (Saudi Arabia).

Also, the correlations between the dependent and independent variables (i.e., the attitudes of the teachers and their ages, teaching experience and speaking assessment experience) in this PhD research contribute important findings to the literature because of the following: (a) the correlation between speaking assessment experience with the EFL teachers' attitudes to intelligibility and

nativeness has not been reported in previous research, including language attitudes research in general; and (b) previous literature has reported correlation between teaching experience and age with language attitudes in general only; it has not addressed the correlation between teaching experience and age with specifically EFL teachers' attitudes to intelligibility and nativeness.

7.2 Brief Answers to the Study's Research Question

The research question set for this exploratory PhD study is: what evidence is there that the attitudes of the NS and NNS EFL teachers in Saudi Arabia are oriented towards nativeness or intelligibility when assessing pronunciation? This main research question has the following research sub-questions:

a. Are the NS and NNS EFL teachers oriented to intelligibility or nativeness when doing an assessment of students' speech task? Are the two groups of teachers different in this regard?
b. Are the NS and NNS EFL teachers oriented to intelligibility or nativeness when doing an acceptability task? Are the two groups of teachers different in this regard?
c. Do they demonstrate attitudes towards intelligibility or nativeness in an online questionnaire? Are the two groups of teachers different in this regard?
d. Are their attitudes in the practical tasks and the online questionnaire in harmony or in conflict? If in conflict, why?
e. Are their attitudes in the practical tasks and the online questionnaire affected by any factors such as any personal background information (e.g., age and experience)? Or such as the institutional assessment policies?

These sub-questions are answered briefly in the following subsections. Greater elaboration and discussion are given in the later subsections.

7.2.1 First sub-Question: Attitudes to Intelligibility and Nativeness in Assessment of Students' Speech Task

In the assessment of students' speech task, none of the differences between the attitudes to intelligibility and nativeness within and across the two groups of teachers were statistically significant ($p > 0.05$), and all the effect sizes were negligible. Thus, it seems that neither group of

teachers (NNS and NS EFL teachers) was more oriented to intelligibility or nativeness in the acceptability task.

7.2.2 Second sub-Question: Attitudes to Intelligibility and Nativeness in Acceptability Task

Like the assessment of students' speech task, none of the differences between the attitudes to intelligibility and nativeness within and across the two groups of teachers in the acceptability task were statistically significant, and all the effect sizes were negligible. Thus, it seems that neither group of teachers was more oriented to intelligibility or nativeness.

7.2.3 Third sub-Question: Attitudes to Intelligibility and Nativeness in Online Questionnaire

The findings of the online questionnaire revealed that there was not a statistically significant difference between the attitudes of the NNS EFL teachers' group to both intelligibility and nativeness, and the effect size was negligible. Thus, this group failed to show more orientation to intelligibility or nativeness. However, there was a statistically significant difference between the attitudes of the NS EFL teachers' group to both intelligibility and nativeness, and the effect size was medium. Thus, this group were more oriented to intelligibility than nativeness. The findings also showed that the NS EFL teachers were more oriented to intelligibility than the NNS EFL teachers in the online questionnaire as the difference between the two groups in this regard was statistically significant and the effect size was medium. However, the attitudes of both groups were similar as regards nativeness since the difference between them in this regard was not statistically significant and the effect size was negligible.

7.2.4 Fourth sub-Question: Correlations between Attitudes in Practical Tasks and Online Questionnaire and Reasons behind Lack of Correlations

The findings of the correlation analyses showed that none of the correlations between the attitudes of both groups of teachers in the practical tasks (i.e., the assessment of students' speech task and the acceptability task) and in the online questionnaire were statistically significant ($p >$ .05) and, where there were correlations, these correlations were weak (r = .10 to .29). Thus, it seems that the participants' attitudes in the online questionnaire were not reflected in their behaviour in the practical tasks.

Interviews with the selected six participants revealed the following for the lack of correlations: five participants were not sure about the reason behind their attitudes, whereas the sixth remaining participant did not answer the question directly making it difficult to draw a conclusion. However, based on the teachers' behaviour and the explanations by two of these participants, it seems that the online questionnaire failed to reveal their real attitudes, which is likely to mean that their behaviour in the practical tasks (i.e., the assessment of students' speech task and the acceptability task; which shows no more orientation of attitudes towards intelligibility or nativeness) is better representative of their real beliefs.

7.2.5 Fifth sub-Question: Factors Affecting Participants' Attitudes

As mentioned in section 5.12.7 (Chapter 5), correlation analyses were conducted in order to investigate whether the teachers' attitudes in the study tasks were affected by their ages, teaching experience or speaking assessment experience. None of the correlations (N = 36) between the attitudes of both groups of teachers in the study tasks and their ages, teaching experience and speaking assessment experience were statistically significant. Also, wherever there were correlations, the correlations were weak. Two exceptions to this (out of 36 correlations) were the correlation between the NNS EFL teachers' ages and their attitudes to intelligibility in the online questionnaire and the correlation between the NNS EFL teachers' speaking assessment experience and their attitudes to nativeness in the online questionnaire; these two correlations were positive and fairly strong, and thus they indicate that (a) the older the NNS EFL teachers were, the more oriented to intelligibility they were in the online questionnaire, and (b) the longer experience these teachers had, the more oriented to nativeness they were in the online questionnaire. As a result, it generally seems that age, teaching experience and speaking assessment experience were not associated with the participants' attitudes.

The six participants interviewed were also asked about any other factors affecting their attitudes in the study tasks. Their answers were as follows: one participant was not sure; two referred to their experience as L2 English learners at school; one referred to his experience of learning English in general; and the remaining two referred to the speaking assessment policies at their institution.

7.3 Attitudes of Participants to Intelligibility and Nativeness in Light of Previous Literature

In light of previous literature, the subsections below discuss the findings of RQ1a, RQ1b and RQ1c of the main research question:

- Are the NS and NNS EFL teachers oriented to intelligibility or nativeness when doing two practical assessment tasks? Are the two groups of teachers different in this regard?
- Do they demonstrate attitudes towards intelligibility or nativeness in an online questionnaire? Are the two groups of teachers different in this regard?

7.3.1 Attitudes within Teachers' Groups in Practical Tasks and Online Questionnaire

The findings of the paired-samples t-tests conducted for the data of the practical tasks (i.e., the assessment of students' speech task and the acceptability task) showed that the differences between the data of both groups of teachers were not statistically significant and all but one of the effect sizes were negligible: the differences between the intelligibility and nativeness features detected by both groups of teachers in the assessment of students speech task and the differences between the intelligible and nativelike pronunciations accepted by both groups of teachers in the acceptability task of students speech tasks. Thus, according to the hypotheses of the practical tasks, it seems that the teachers of both groups failed to show more orientation to intelligibility or nativeness in the practical tasks.

The findings of the paired-samples t-tests conducted for the data of the online questionnaire showed that the difference between the means of the NNS EFL teachers' responses to the intelligibility and nativeness statements was not statistically significant and the effect size was negligible. Thus, it seems that the NNS EFL teachers failed to show more orientation to intelligibility or nativeness. However, the findings of the paired-samples t-tests showed that the difference between the means of the NS EFL teachers' responses to the intelligibility and nativeness statements was statistically significant and the effect size was medium; the means of their responses to the intelligibility statements were lower than the means as regards the nativeness statements. Thus, it seems that the NS EFL teachers were clearly more oriented to intelligibility than nativeness.

Given the attitudes of both groups of teachers in the study tasks mentioned above, it seems that the teachers in general did not show strict adherence to nativeness, but they may have been more oriented to intelligibility than nativeness given the NS EFL teachers' clear orientation to

intelligibility over nativeness in the online questionnaire. It seems that the teachers' attitudes in the study tasks do not strictly comply with the strong position that nativeness had held in L2 pronunciation teaching and assessment for a long time (Levis, 2005) and in some empirical research looking at EFL teachers' attitudes to intelligibility and nativeness (e.g., Ahn, 2011; Baker, 2011; Tsui & Bunton, 2000; Sifakis & Sougari, 2005; Stanojević et al., 2012; Young & Walsh, 2010). Some of this research is relatively old, and it is not of surprise to see that it shows orientation of attitudes towards nativeness as this used to be the norm (Cook, 1999). Also, most of this research did not (a) look at the attitudes of the teachers through practical assessment tasks, adopting the LFC as a framework for intelligibility and nativeness, nor (b) look at the attitudes in pronunciation assessment. These different conditions to the current study might have been a reason for the different attitudes in such research and the current study. For example, the teachers in Tsui and Bunton (2000) showed general acceptance of NSs as a source of authority in terms of language use, although the teachers were somewhat different in terms of the level of acceptance of NSs as a source of authority; some showed total acceptance while others did not. Similarly, the majority of the teachers in Sifakis and Sougari (2005) (a) showed approval of their own accents as they believed they sounded nativelike, (b) thought that their students should sound nativelike, (c) exposed their students to only NSs' conversations, and (d) thought that English belonged to NSs only.

Ahn (2011), Baker (2011a), Stanojević et al. (2012), and Young and Walsh (2010) are more recent studies, but still at least 10 years out of date. They all showed orientation of EFL teachers towards nativeness, unlike the current study. Ahn (2011) found orientation of an EFL teacher's attitudes towards nativeness through the teacher's reported beliefs and reported practices. The reported practices of the teacher involved his pronunciation assessment practices, but the reported beliefs did not involve attitudes in pronunciation assessment. Also, the teachers in Baker (2011a) reported (a) they mostly focused on suprasegmentals in their teaching (which this study considers to be mostly features of nativeness; see section 3.5, Chapter 3), (b) they thought suprasegmentals were the most important in teaching, and (c) they expressed their need for more education and training on suprasegmentals. Similarly, Stanojević et al. (ibid.) found that the EFL teachers in their study adopted nativeness as a benchmark for teaching pronunciation based on their reported beliefs in interviews. Finally, the EFL teachers' reported practices and reported beliefs in Young and Walsh (2010) also show somewhat clear orientation of their attitudes towards nativeness. The great majority of the teachers in their study reported their teaching of NS varieties and expressed their interest in teaching NS varieties.

However, the slight orientation of EFL teachers' attitudes in favour of intelligibility over nativeness seen in the current study seems to support the noticeable shift to intelligibility in L2 pronunciation teaching and assessment (Levis, 2005) and in some empirical research looking at EFL teachers' attitudes to intelligibility and nativeness (e.g., Baker & Burri, 2016; Bøhn & Hansen, 2017; Buckingham, 2015; Couper, 2017; Foote's at al., 2012; Sannes, 2013; Timmis (2002). The majority of this research is relatively recent, and it shows a shift of attitudes from nativeness towards intelligibility which can also be observed in the current study. For example, the EFL teachers in Baker and Burri (2016) highlighted the importance of feedback to improve intelligibility in pronunciation. They also highlight intelligibility as a purpose of pronunciation teaching. Bøhn and Hansen (2017) also found that the EFL teachers in their study were oriented to intelligibility when interviewed and doing a Likert scale questionnaire. In the interview and the Likert scale questionnaire, the teachers showed interest in intelligibility over nativeness. In the Likert scale in particular, the teachers gave importance to pronunciation features which contribute more to intelligibility, such as segments rather than suprasegmentals. In Buckingham (2015), the majority of the EFL teachers in that study put emphasis on the importance of clear English rather than nativeness. They also expressed approval of their own L2 accents. The EFL teachers in Couper (2017) also reported that effective communication and intelligibility were the goals of their pronunciation teaching rather than nativeness and reported their focus on segments in pronunciation teaching and on making their students proud of their L2 accents and their identities. Similarly, in an online survey used in Foote et al. (2012), the teachers' responses regarding attaining nativeness in pronunciation and foreign accents in teaching pronunciation show that the teachers' attitudes were more oriented to nativeness than intelligibility. Finally, the EFL teachers investigated by Sannes (2013) put more emphasis on communication and ELF rather than nativeness. Although the findings of such research concur with the findings of the current study, they did not consider (a) pronunciation assessment (except for Bøhn and Hansen, 2017), (b) the LFC as a framework for intelligibility and nativeness, nor (c) actual assessment practices when investigating the teachers' attitudes. Given this, the current study makes an important contribution to the development of intelligibility as a priority in ELT and assessment.

In a much earlier study focusing on attitudes, whose findings seem to be an exception to the norm in earlier research, Timmis (2002) found that the attitudes of the majority of the EFL teachers in their research were more oriented towards intelligibility than nativeness. The conditions in Timmis (ibid.) are different to those in the current study; Timmis (ibid.) used a survey to explore the teachers' attitudes and looked at the attitudes of the teachers in pronunciation in

general rather than pronunciation assessment in specific like the current study. Also, he did not consider the LFC as a framework to differentiate between intelligibility and nativeness like the current study. However, we can still consider Timmis' study as an early example of work indicating the developing perceived importance of intelligibility over nativeness among ELT teachers, which the current study expands on to include orientation towards intelligibility in ELT assessment practice.

Given the orientation of the attitudes of both groups of teachers in the current study tasks slightly more towards intelligibility than nativeness, it seems that the teachers do not necessarily perceive nativeness as a measure of success in L2 acquisition, as it has been traditionally seen (Angelovska & Hahn, 2009; Cook, 1999; Deterding, 2010). Also, the teachers seem to reject to some extent the Chomskyian notion that suggests the native speaker is the leading authority on L2 language use (McKay, 2012). They also seem to somewhat reject SLA research's conventional position that success and failure in L2 is decided based on native speaker competence (Bley-Vroman, 1983).

7.3.2 Attitudes among Individual Teachers in Practical Tasks and Online Questionnaire

On an individual basis, the teachers in the current study behaved similarly in terms of intelligibility and nativeness in the following cases: (a) both groups of teachers in the assessment of students' speech task, (b) the NS group in the acceptability task, and (c) the NNS group in the online questionnaire. In these cases, the outputs of the paired-samples t-tests conducted for the data of the study tasks show that the standard deviations for the nativeness scores are higher than the standard deviations for the intelligibility scores. This indicates that there were greater variations in the attitudes of the individual participants towards nativeness than there were for intelligibility. The participants seem to be only slightly different from each other regarding their attitudes towards intelligibility in those cases, compared to their attitudes towards nativeness where the participants were more different from each other. This suggests that the extremes of orientation to nativeness seem to be greater than the extremes of orientation to intelligibility. On an individual basis, this may indicate that the ideology of native-speakerism is still stronger in the mindset of some teachers than intelligibility, while some other teachers totally reject this ideology. However, the NNS group in the acceptability task and the NS group in the online questionnaire behaved differently on an individual basis. The outputs of the paired-samples t-tests conducted for the data of the acceptability task and the online questionnaire show that the standard deviations of the intelligibility scores for those two

cases are higher than the standard deviations of the nativeness scores. This suggests that the teachers in those two cases were more different from each other regarding their attitudes towards intelligibility than they were regarding their attitudes towards nativeness.

The variations between the individual teachers in the majority of cases mentioned above are similar to a case in Bøhn and Hansen (2017). In a part of their study, Bøhn and Hansen (ibid.) investigated EFL teachers' attitudes towards intelligibility and nativeness in pronunciation assessment using interviews. The findings of these interviews showed significant variations among the interviewees in terms of their attitudes towards nativeness: five (out of 21) interviewees did not find nativeness important at all, five found it slightly important, five found it somewhat important, and six found it extremely important. Although the variations among the teachers in the majority of cases in this study agree with the ones in Bøhn and Hansen (ibid.), most of the variations in the current study were found in practical tasks adopting the LFC as a framework for intelligibility and nativeness, while the variations in Bøhn and Hansen (ibid.) were identified in reported beliefs in interviews only. Also, no previous research has looked into variations between individual participants in terms of their attitudes towards intelligibility (as far as the author of this study is aware), like the current study. Thus, the variations among individual teachers regarding their attitudes towards intelligibility and nativeness in the current study and the explanations of these variations mentioned above make an important contribution to our understanding of the development of intelligibility as a priority over nativeness in teachers' assessment practices as well as in their attitudes.

7.3.3 Special Case of Teachers' Attitudes

The attitudes of the NS EFL teachers in the online questionnaire in particular is a unique case, as noted earlier in section 7.3.1. Only in this case did the statistical analysis show a significant difference indicating the NS EFL teachers' attitudes were clearly orientated to intelligibility over nativeness in the online questionnaire. This is different to the case of the NNS EFL teachers in the same questionnaire and the cases of both groups of teachers (i.e., the NNS and NS EFL teachers) in the assessment of students' speech task and the acceptability task. It is possible that the attitudes of the NS EFL teachers were in fact significantly oriented to intelligibility in the assessment of students' speech task and the acceptability task like the case in the online questionnaire, but intelligibility was not measured appropriately in the assessment of students' speech task and the acceptability task; this could be a limitation of the study. It might be the case that the framework

used for intelligibility (i.e., the LFC) in the assessment of students' speech task and the acceptability task was not an appropriate one. In the online questionnaire, attitudes to intelligibility were measured based on statements expressing positions regarding intelligibility in general, i.e., "Students who are intelligible (understandable) will get higher marks". In the assessment of students' speech task and the acceptability task, the LFC was used for measuring attitudes to intelligibility; however, the LFC has been criticised and not all the empirical studies investigating the LFC have supported its proposals (see sections 3.5.1 and 3.5.3, Chapter 3). Alternatively, the case might be that the assessment of students' speech task and the acceptability task were not built appropriately to measure attitudes to intelligibility. As mentioned in sections 5.4.1.1 and 5.4.3 (Chapter 5), the assessment of students' speech task and the acceptability task were developed by the author of this study as there is no suitable instrument in the literature that can be adopted or adapted to measure attitudes to intelligibility in pronunciation assessment. Since these two tasks were developed almost from scratch, they should be subject to problems. Thus, the possibility that there are problems with the assessment of students' speech task and the acceptability task might have led to inaccurate conclusions regarding the NS EFL teachers' orientation towards intelligibility.

The clear orientation of the NS EFL teachers' attitudes towards intelligibility in the online questionnaire in contrast to their attitudes in the assessment of students' speech task and the acceptability task may give support to Seidlhofer's (2011) claim regarding the attitudes towards nativeness. Seidlhofer (ibid.) argued that the belief that nativeness is more important than intelligibility has changed, whereas practices in this regard have not. This stands true in the case of the attitudes of the NS EFL teachers, where their attitudes were significantly oriented to intelligibility rather than nativeness in their reported attitudes in the online questionnaire, while their behaviour in the assessment of students' speech task and the acceptability task failed to reveal more orientation to intelligibility or nativeness.

7.3.4 Attitudes across Teachers' Groups in Practical Tasks and Online Questionnaire

The findings of the independent-samples t-tests showed that the differences between the two groups of teachers in the practical tasks (i.e., the assessment of students' speech task and the acceptability task) were not statistically significant and the effect sizes were negligible in one case, small in two cases, and medium in one case: the differences between the two groups regarding the intelligibility and nativeness features detected in the assessment of students speech task and regarding the intelligible and nativelike pronunciations accepted in the acceptability task. Thus, it

may be concluded that the teachers of both groups were in general similar in their attitudes in the practical tasks.

The findings of the independent-samples t-tests conducted for the data of the online questionnaire showed that the differences between the two groups of teachers regarding the means of their responses to the nativeness statements were not statistically significant and the effect size was negligible. However, the differences between the two groups of teachers regarding the means of their responses to the intelligibility statements were statistically significant and the effect size was medium; the means of the NS EFL teachers' responses to the intelligibility statements were lower than the NNS EFL teachers' degrees. Thus, it seems that the attitudes of both groups of teachers were similar regarding nativeness: both groups of teachers showed almost equal orientation to nativeness. However, regarding intelligibility, the NS EFL teachers' group seemed to be more oriented to intelligibility compared to the NNS EFL teachers' group.

If the attitudes across both groups of teachers in the practical tasks (i.e., the assessment of students' speech task and the acceptability task) in this PhD study are considered similar regarding intelligibility, this may support the case in Buckingham (2015) mentioned above. The majority of the teachers in Buckingham's study, who were both NNS and NS EFL teachers, adopted an intelligibility position regarding their practices. In this regard, Buckingham (ibid.) stated that there was only a little difference between the NNS and NS EFL teachers.

The attitudes across both groups of teachers in the online questionnaire only in this study regarding intelligibility which showed more orientation of the NS EFL teachers to intelligibility than the NNS EFL teachers may offer support to the attitudes found in Timmis (2002, as mentioned above). Timmis (ibid.) found that most of the EFL teachers in his study, especially NS teachers, show more orientation to intelligibility than nativeness. 37% of the NNS teachers in his study preferred intelligibility, compared to 32% of the NNS teachers who preferred nativeness, and 41% of the NS teachers preferred intelligibility, compared to 20%. However, given the attitudes within both groups of teachers in this PhD study and across the two groups in all the study tasks, it seems that the gap between the NNS and NS EFL teachers is closing since the attitudes of both groups have demonstrated a shift of orientation from nativeness towards intelligibility.

The conclusions described above regarding the differences between both groups of teachers (i.e., NNS and NS EFL teachers) in terms of their attitudes in the study tasks can be further

justified by the attitudes within groups. As mentioned in the previous section above, the findings of the current study showed that the attitudes of both groups of teachers failed to show more orientation to intelligibility or nativeness in the following cases: the NNS EFL teachers in the assessment of students' speech task, the acceptability task and the online questionnaire; and the NS EFL teachers in the assessment of students' speech task and the acceptability task. Similarly, when comparing the two groups of teachers regarding their attitudes to intelligibility and nativeness in the assessment of students' speech task and the acceptability task, none of them were more oriented to intelligibility or nativeness than the other group. In the online questionnaire, the attitudes of the NS EFL teachers were more oriented to intelligibility than nativeness as mentioned in the previous section, while the attitudes of the NNS EFL teachers failed to show more orientation to intelligibility or nativeness. Thus, when comparing both groups of teachers in terms of their attitudes in the online questionnaire, the NS EFL teachers were more oriented to intelligibility than the NNS EFL teachers.

Since there were no significant differences between the two groups of teachers (i.e., the NNS and NS EFL teachers) in terms of their attitudes to intelligibility and nativeness in most parts of the study tasks as mentioned earlier in this section, it seems that speaking English as a native language is no longer a priority. Comparison between the two groups of teachers in this study was conducted as it is a common practice to compare between NNSs and NSs on different aspects in Applied Linguistics empirical research (see, e.g., Dewaele, 2018; Morris-Adams, 2016; Nymeyer et al., 2022), as far as the author of this study is aware. Also, it was assumed by the author of this study that the two groups would have different attitudes regarding intelligibility and nativeness. It was assumed by the author that the NS group would prefer intelligibility as they come into contact with many NNSs, and the intelligibility of those speakers to them are more important than being nativelike. However, it was assumed by the author of this study that the NNS group would prefer nativeness as the materials and sources those speakers used to learn English and they are using to teach English and assess English proficiency should have been centred on nativeness. All these assumptions seem to be incorrect to a large extent based on the findings of comparison between the NNS and NS groups in this study.

7.4 Correlations between Participants' Attitudes in Practical Tasks and Online Questionnaire in Light of Previous Literature

In the light of previous literature, this section discusses the findings of RQ1c:

> Are the NNS and NS EFL teachers' attitudes in the practical tasks and the online questionnaire in harmony or in conflict? If in conflict, why?

The findings of the correlation analyses conducted for the data of both groups of teachers showed that there were no correlations between the dependent variables in the practical tasks (i.e., the assessment of students' speech task and the acceptability task) and the online questionnaire in half of the cases, while the other half showed weak correlations (according to the correlation coefficient). Thus, this may indicate that in general the attitudes of both groups of teachers in the online questionnaire were not reflected in their behaviour in the practical tasks. Also, none of the correlations was statistically significant according to the p-value analyses.

The findings of the correlation analyses seem to be consistent with research on other topics which found that reported attitudes are not necessarily correlated with behaviour (e.g., Krezanoski et al., 2018; Pantos & Perkins, 2013). For example, Pantos and Perkins (2013) investigated implicit and explicit attitudes towards foreign and US accented speech. The implicit attitudes were measured using participants' reactions to audio stimuli, and the explicit attitudes were measured using self-report questionnaires. The participants exhibited accent bias in their reactions to the audio stimuli, while they did not exhibit the same bias in the self-report questionnaires.

The findings of the correlation analyses seem to be inconsistent with the empirical research that show the attitudes of EFL teachers towards intelligibility and nativeness are similar across their reported beliefs and reported practices (e.g., Ahn, 2016; Baker, 2011). For example, Ahn (2011) found similarity of orientation of an EFL teacher's attitudes towards nativeness through the teacher's reported beliefs and reported practices. The reported practices of the teacher involved his pronunciation assessment practices, but the reported beliefs did not consider attitudes in pronunciation assessment. Also, based on EFL teachers' reported practices in Baker (2011a), those teachers focused on nativeness when teaching as they mostly focused on suprasegmentals (which are nativeness features; see section 3.5, Chapter 3) in their teaching. In their reported beliefs, the EFL teachers in Baker (ibid) thought suprasegmentals were the most important in teaching and

expressed their need for more education and training on these features. Both Ahn (2016) and Baker (2011a) depended only on reported practices rather than actual practices, unlike the current study. Also, they did not adopt the LFC as a framework for intelligibility and nativeness like the current study, nor did they test all the attitudes of the EFL teachers in pronunciation assessment. However, the findings of the correlation analyses may give support to the empirical research that has found that EFL teachers' beliefs regarding intelligibility and nativeness are not correlated with their orientation in reported practices (e.g., Buss, 2016). The teachers' reported beliefs in Buss's (2016) study showed more orientation to intelligibility than nativeness as the teachers showed more interest in intelligibility when asked about the goals of pronunciation teaching and the models for teaching pronunciation. However, the teachers' reported practices were more oriented to nativeness as the teachers reported their focus on phonological and phonetic features that are nativeness features (see section 3.5, Chapter 3), such as dental fricatives. Like Ahn (2016) and Baker (2011a) mentioned above, Buss (2016) depended only on reported practices rather than actual practices, unlike the current study. Buss (ibid.) neither looked at the attitudes in pronunciation assessment, nor did he consider the LFC as a framework for intelligibility and nativeness. Thus, the comparison between the current study and Buss (ibid.) is not without problems.

The non-correlation between the teachers' attitudes in the online questionnaire and the practical tasks (i.e., the assessment of students' speech task and the acceptability task) may be attributed to some reasons mentioned in previous literature. For example, perceptions are often hidden behind a social façade (Dalton-Puffer et al., 1995; Giles & Billing, 2004), and thus somebody may express an attitude, but his/her behaviour contradicts that attitude. Similarly, the non-correlation may be a result of the teachers trying to give socially acceptable responses in the online questionnaire (Lenski & Leggett, 1960), which were not their real beliefs as their behaviour in the online questionnaire were not reflected in the practical tasks. It may also be attributed to difficulty in measuring language attitudes (Allport, 1935, as cited in Hsu, 2012). They may also be explained by teachers' awareness of being assessors and thus they did not allow their own underlying assessment criterion (e.g., nativeness) to be activated and influence their assessment behaviour (Ajzen & Fishbein, 2005, p.182). Another possible explanation for the non-correlation between the teachers' attitudes in the online questionnaire and the practical tasks is that the non-correlation may be a result of the teachers' previous training or education to reduce subjectivity in scoring decisions (Lumley & McNamara, 1995; McNamara, 1996; Weigle, 1998); thus the teachers may have tried to avoid their own beliefs during the practical tasks and thus they based their responses on other criteria such as policies adopted by their institutions. Another possibility is that the teachers'

attitudes towards intelligibility and nativeness were not activated while doing the practical tasks (Fazio, 1986, 1990, 1995; Fazio & Towles-Schwen, 1999), possibly because of low motivation for doing the study tasks (Hsu, 2012); this in return may have resulted in the inconsistency between the attitudes in the online questionnaire and the practical tasks. Further possible reasons for the non-correlation between the teachers' attitudes in the online questionnaire and their behaviour in the practical tasks are the limitations of the methodology for collecting the data mentioned in sections 7.6 below and 8.4, Chapter 8.

7.5 Effects of Participants' Backgrounds on Attitudes in Light of Previous Literature

In light of previous literature, this section discusses the findings of RQ1d:

> Are the NNS and NS EFL teachers' attitudes in the study tasks affected by any factors such as any personal background information (i.e., age and experience)? Or such as institutional assessment policies?

As noted in section 6.5 (Chapter 6), none of the correlations between the attitudes of both groups of teachers (i.e., the NNS and NS EFL teachers) in the study tasks and the independent variables were statistically significant and that the great majority of cases indicated non-correlation or weak correlation. This seems to be inconsistent with the research that has found correlation between teaching experience and/or age and language attitudes (e.g., Agathopoulou, 2010; Arrieta, 2016; El-Dash & Busnardo, 2001; Hsu, 2012; Libben & Rossman-Benjamin, 1992; Richards, 1996; Richards et al., 1992). For example, Arrieta (2016) investigated the English varieties EFL teachers wanted to teach. The attitudes of the teachers in Arrieta (2016) failed to show more orientation to intelligibility or nativeness. They showed some agreement with the importance of English NNS varieties and the importance of their instruction. The teachers also confirmed teaching these varieties only from time to time. However, Arrieta (ibid.) found that the more experienced and older teachers were more oriented in their attitudes to nativeness than the teachers who are younger and less experienced. Based on this, Arrieta (ibid.) referred to age and teaching experience as possible effects on the EFL teachers' language attitudes. This was not, however, reflected in the current study. The conditions in Arrieta (ibid.) and the studies mentioned above are different to the ones in the current study; none of these studies looked into EFL teachers' reported attitudes and actual practices in pronunciation assessment while considering the LFC as a framework for intelligibility and nativeness. Thus, this could be the reason between the inconsistency between the current study

and these studies. This could also be attributed to the limitations of the methodology used to collect the data of this study mentioned in sections 7.6 and 8.4 (Chapter 8). Given this, the current study makes an important contribution to our understanding of the possible factors affecting attitudes towards intelligibility and nativeness in pronunciation assessment.

In the follow-up interview, experience as L2 English learners was reported by some participants as a possible effect that made their attitudes in the study tasks oriented towards nativeness. Those participants reported that the materials they used for learning at school or for self-learning were based on native speakers' norms, and thus this may have affected their attitudes towards nativeness. This result is consistent with the findings of some previous research (e.g., Baker, 2011; Borg, 2003; Libben & Rossman-Benjamin, 1992). However, the conditions that are different in the current study and such research like investigating attitudes in pronunciation assessment may make the comparison between the current study and such research somewhat invalid. For example, Baker (2011) investigated the development of EFL teachers' attitudes regarding pronunciation teaching. She used interviews and classroom observations to collect the data of her study. She found that the teachers seemed to be influenced by their previous education including L2 English learning. This reflected on their attitudes regarding pronunciation teaching and their practices in the classroom. For example, one of the teachers focused on teaching word stress in the classroom as word stress used to be emphasised in the materials she had studied when she was at school. Since word stress is a nativeness feature (see section 3.5, Chapter 3), her attitudes in this regard were oriented towards nativeness as she focused on teaching it. Thus, it might be the case that experience as L2 learners was associated with the attitudes of some participants in the current study.

Two participants in the follow-up interview thought that speaking assessment policies at their institution had an effect on their attitudes in the tasks in this study. The effect of speaking assessment policies on language attitudes has not been investigated by previous research (as far as the author of this study is aware), and thus any conclusion regarding the association between speaking assessment policies and attitudes to intelligibility and nativeness in the current study needs to be validated with more investigation.

7.6 Limitations of Study

There are several limitations of this PhD study, which may reduce the reliability of the findings. Firstly, the study included EFL teachers drawn mainly from only KAU (see section 5.5, Chapter 5). The reason that the majority came from this university is that they were the researcher's colleagues, and thus it was convenient to approach them, and access other relevant materials they used such as the assessment policy. Thus, had the EFL teachers been selected from several different universities in Saudi Arabia, the study would have better represented EFL teachers in Saudi Arabia.

Another limitation of the study is recruiting few female EFL teachers besides male teachers; male teachers were the great majority. The plan was to recruit only male EFL teachers to avoid gender as a possible variable affecting results, but obtaining a statistically enough NS sample with male teachers only was not possible. Having a statistically enough sample is important to be able to carry out inferential statistics. Access to study participants was restricted due to the timing of data collection. The first occasion when the author of this study travelled to Saudi Arabia to collect the data coincided with summertime when the majority of teachers were on holiday. The second time when the author travelled to Saudi Arabia again to collect more data coincided with the COVID-19 pandemic. Thus, the author had to contact prospective male participants after the first and second time of travelling to Saudi Arabia to collect the data via email, but still not statistically enough male NS participants were reached, although hundreds from KAU and other Saudi universities were emailed. Thus, female NS EFL teachers from KAU were contacted via email to ask them for participation in the study. As a result, a few female teachers participated in the study. The study should have recruited only statistically enough males for both the NNS and NS groups so as to avoid any additional or accidental effect from the additional variable of gender.

One further limitation of the study is that the number of both NNS and NS EFL teachers in the study were not equal (see section 5.5, Chapter 5); there were 33 NNS EFL teachers and 24 NS EFL teachers. This is because further recruitment of participants was restricted by the reasons mentioned above, although the plan was to recruit at least 30 NNS teachers and 30 NS teachers.

Two more limitations are related to the assessment of students' speech task on its own. The recordings of the students in the task were made using a mobile phone; a teacher did this on behalf of the author of this study (see section 5.4.1.3, Chapter 5). Thus, this may have resulted in low quality recordings which may have led to poor reflection of the participants' attitudes; the

participants might have detected features as problematic, not because they thought they were problematic, but because they misheard these features and thus detected them as problematic. Thus, the recordings should have been made using professional recording equipment instead. That was not possible because the author of this study could not meet the students face to face as he was in UK while making the recordings and the students were at KAU, Saudi Arabia. Also, there was no recording equipment available at KAU at the time of data collection.

Also, the proficiency levels of the students recorded in the assessment of students' speech task were decided by the teacher who helped collect the recordings on behalf of the researcher based on his familiarity with the students rather than conducting placement tests to put the students into different levels of proficiency (see section 5.4.1.2, Chapter 5). This teacher selected three students from three different proficiency levels: i.e., advanced, intermediate and beginner students. This may have resulted in recordings of the free speech task that did not represent all the phonological and phonetic features listed in the assessment of students' speech task. As mentioned in section 5.4.1.2 (Chapter 5), there were two speaking tasks that the learners performed. One of them is reading a script loudly which was based on the intelligibility and nativeness features (i.e., the core and non-core features of the LFC; see section 5.4.1.1, Chapter 5), and the other was a free speech task, where all the intelligibility and nativeness features may not emerge. In the view of the author of this study, having L2 English learners with different levels of proficiency doing the free speech task may have helped produce recording samples of this task that include different intelligibility and nativeness features, but which jointly represented all the intelligibility and nativeness features. Thus, the proficiency levels of these students should have been determined using a better way than the teacher's familiarity with the students such as a proficiency test.

One final limitation concerns the data collection. The data for the study were collected online rather than through face-to-face meetings with the participants (i.e., the EFL teachers in the study; see section 5.7, Chapter 5). The original plan for data collection was to meet face-to-face with the participants and handle all the data collection procedure personally. Unfortunately, the time the researcher travelled to the context of the study to collect the data there were only a few participants available. Following this, three teachers were asked to collect the data on behalf of the researcher, but the contact with two of them could not be maintained, and the third teacher expressed his inability to handle it alone. Following this, physical access to the participants was restricted due to the COVID-19 pandemic. Thus, a better control over the data collection procedures can be maintained through physical collection of the data. Collecting data physically may help provide

more accurate data through controlling timing and environment, handling any queries the participants raise regarding the study instruments, and so on.

7.7 Summary of Chapter

This chapter has discussed the findings in relation to the research questions in light of previous literature. It has firstly discussed the attitudes of the EFL teachers study tasks towards nativeness and intelligibility. Next, the chapter has discussed the correlations between the teachers' attitudes in the practical tasks (i.e., the assessment of students' speech task and the acceptability task) and the online questionnaire. Then, it has discussed the effect of the independent variables and any other factors on the teachers' attitudes. Following this, the chapter has discussed the other findings that have revealed following in-depth analysis of the data. Finally, the chapter concluded with the limitations of the study.

CHAPTER 8: CONCLUSION

8.1 Introduction

This study has investigated the attitudes of NNS and NS EFL teachers' attitudes to intelligibility and nativeness in pronunciation assessment. It made use of direct and indirect elicitation methods to find out the attitudes.

This concluding chapter summarises the major findings of this study. It then presents the contributions this study made to the literature. It then introduces pedagogical implications based on the current study. It finally presents recommendations based on the limitations of this study.

8.2 Contributions of Study

The main contribution of the study is to demonstrate that, as observed in contemporary research, EFL teachers at a tertiary institution in Saudi Arabia are showing almost equal interest in nativeness and intelligibility in pronunciation assessment. This contrasts the situation that used to exist in the past when nativeness was given more importance than intelligibility. As discussed in the previous chapters, nativeness used to be clearly favoured by L2 researchers, teachers and learners over intelligibility (Levis, 2005). The same conclusion was also reached by some research investigating EFL teachers' attitudes towards intelligibility and nativeness, which is at least 10 years out of date (e.g., Ahn, 2011; Baker, 2011; Tsui & Bunton, 2000). Another contribution is to show the orientation of EFL teachers' attitudes towards intelligibility in pronunciation assessment, unlike the majority of the empirical research that has found the same results but has not looked specifically at pronunciation assessment (e.g., Arrieta, 2016; Baker & Burri, 2016; Buckingham, 2015). One further contribution is to show that NS EFL teachers are more oriented to intelligibility in their reported attitudes than NNS EFL teachers, unlike the practices of both groups of teachers which showed almost equal orientation to both intelligibility and nativeness.

One further contribution is to show that reported beliefs are not necessarily reflected in pronunciation assessment practices. This was indicated by the lack of correlation between the EFL teachers' attitudes in the online questionnaire and the practical assessment tasks (i.e., the assessment of students' speech task and the acceptability task), which was revealed by the correlation analysis between the teachers' responses to the online questionnaire and their

responses to the practical tasks. One last contribution is to show that age, teaching experience and speaking assessment experience may not have an effect on attitudes towards intelligibility and nativeness in pronunciation assessment. This was indicated by the lack of correlation between these variables and the EFL teachers' pronunciation assessment practices in this study. However, L2 learning experience could be associated with attitudes towards intelligibility and nativeness in pronunciation assessment as it was reported by some teachers in the follow-up interview as a possible effect on the orientation of their attitudes towards nativeness in the pronunciation assessment tasks in this study.

8.3 Pedagogical Implications

Based on the findings of this research, several possible pedagogical implications arise. For example, given the relevance of intelligibility (particularly Jenkins', 2000, model for intelligibility—the LFC) demonstrated in Chapter 2 and Chapter 3, and since the findings of the EFL teachers' attitudes in this study still showed some orientation to nativeness, it is important that phonological features essential for intelligibility are made clear to EFL teachers. Further training and education for EFL teachers are required where they are oriented towards nativeness so that the teachers understand the importance of intelligible communication in English as a global language and change their attitudes towards intelligibility. Their practices need to be observed to ensure that their beliefs are reflected in their practices, as the EFL teachers' reported beliefs in this study did not seem to be reflected in their practices. The training and education need to be provided to all EFL teachers regardless of their L1 background, as both groups of teachers in this study (i.e., the NS and NNS EFL teachers) showed similar attitudes towards intelligibility and nativeness, where both groups showed almost equal interest in both intelligibility and nativeness when assessing pronunciation. Also, both education and training should be provided to all EFL teachers regardless of their ages and experience since this study showed that age, teaching experience and speaking assessment experience were not associated with EFL teachers' attitudes to intelligibility and nativeness. For a more effective way, institutional teaching and assessment policies need to be amended to reflect the shift towards intelligibility as they seem to be associated with EFL teachers' attitudes towards intelligibility and nativeness, as indicated by the follow-up interviews in this study.

As mentioned in Chapter 3, the LFC is a temporarily working framework for intelligibility. Until another framework for intelligibility has been established in empirical studies, the LFC may be recommended to EFL teachers as a guideline for the phonological and phonetic features essential

for intelligibility in English as a global langauge. However, EFL teachers need to be educated about the core and non-core phonological and phonetic features of the LFC so that they focus on the features important for intelligibility (i.e., the core feature) in their instruction and assessment. The education as regards this is of particular importance given that none of the EFL teachers in this study reported their familiarity with or knowledge of the LFC. This does not mean that learners should be prevented from pursuing nativelike norms if they wish but that the focus should be on enabling them to communicate intelligibly in English for international communication (Walker et al. 2021).

Given the recommendations for adopting an ELF approach to intelligibility stated in Chapter 3, EFL teachers should be aware that English is used in various contexts as a medium for communication rather than as an inflexible knowledge owned by native speakers (Jenkins, 1988). In particular, EFL teachers should recognize that L2 English learners are multi-competent language users (Cook, 1999) or intercultural speakers (Kramsch, 1998) rather than defective communicators (Firth & Wagner, 1997). This study goes some way to demonstrate English language teachers from NS or NNS backgrounds are either aware, or becoming aware, of this and that it is reflected in their attitudes to English pronunciation and in their subsequent assessment behaviour.

8.4 Recommendations

Given the limitations of the study mentioned in Chapter 7 and the limited research into EFL teachers' attitudes, particularly in pronunciation assessment, EFL teachers' attitudes in pronunciation assessment need to be reconsidered in empirical research. The findings of the current study may serve only as a starting point for more detailed investigations into EFL teachers' attitudes in pronunciation assessment. This study has just scratched the surface of a partially explored area of research in EFL teachers' attitudes and in only one context—Saudi Arabia.

Further research into EFL teachers' attitudes towards intelligibility and nativeness in pronunciation assessment should consider these attitudes in different contexts other than Saudi Arabia. They may consider looking at these attitudes in Saudi Arabia, but in different contexts to the one in this study. This study has investigated specifically the EFL teachers at KAU; further research should look at other campuses in Saudi Arabia and in other contexts internationally. Further research may also consider genders separately. In the current study, the majority of the participants are males with only a few female participants; owing to the low number of female participants and

lack of balance in numbers, it was decided not to look at the two genders separately when looking at their attitudes and comparing between the groups. Further research may look at female EFL teachers' attitudes only or seek to have large enough numbers of participants of both genders to be able to make reliable comparisons between female and male EFL teachers' attitudes. There should be enough female and male EFL teachers in the study to be able to carry out inferential statistics if a quantitative research paradigm is adopted.

In the current study, NS and NNS EFL teachers were recruited and their attitudes were investigated and compared. Deciding whether a participant is a NS or NNS was made based on the countries from where each participant comes. Further research may try to find better methods for classifying participants as NSs or NNSs. Such research may need to find out more details about participants' backgrounds before deciding if they are NSs or NNSs. For example, they may ask the participants background questions about the first languages of their parents and where they have spent their childhood. Such questions may help better decide if a participant is a NS or NNS, although classifying the speakers of English as NNSs and NSs is a thorny issue (Davies, 2003; Cook, 2007) and so future research may seek to use a different distinction, such as Jenkins's (2000) Bilingual English Speaker and Non-Bilingual English Speaker.

This study has looked at EFL teachers' attitudes to find out if further education and training for these teachers are required in order that they change their attitudes towards intelligibility (see section 1.2, Chapter 1). Further research could expand into other teachers or professional groups to look at, for example, the attitudes of curriculum and assessment designers. Such persons may still need to change their attitudes towards intelligibility as these attitudes and the practices they subsequently adopt may have an effect on the shift of attitudes towards intelligibility. Curriculum and assessment designers may have an effect on the attitudes of L2 teachers and learners through the teaching and assessment materials they design. Thus, identifying the attitudes of these persons and then providing education and training for them if their attitudes were oriented towards nativeness can help shift their attitudes and, as a result, the attitudes of L2 teachers and learners towards intelligibility.

This study adopted the LFC as a framework for what it is an intelligible feature or a nativelike feature since the LFC is a temporary working model for intelligibility (Isaacs, 2018; Walker, 2010). However, there are several criticisms against the LFC, and there is some empirical research that has found results contradicting the proposals of the LFC. Thus, further research into

attitudes towards intelligibility and nativeness should first test the proposals of the LFC among different learner/speaker groups. Based on the results of this testing, such research should adapt the LFC before using it as a framework that differentiates between intelligibility and nativeness phonetic and phonological features (Zoghbor, 2018).

As discussed above, some previous research has found correlation between age and experience with attitudes, while the current study did not find such correlation. The lack of correlation in this study may be attributed to shortcomings in the research instruments, so further research into this is required so that better understanding of the association between age and experience with attitudes can be established.

Speaking assessment experience was investigated in this study to see if it is correlated with attitudes, and the findings did not reveal that there was any correlation. Previous research did not touch on this aspect of investigation, and thus further research into this is vital to obtain a better understanding of language attitudes interaction with experience among teacher-assessors and provide education and training to assessment designers accordingly. However, in the follow-up interviews presented here, assessment policies were reported as a factor affecting attitudes; thus empirical research is required to confirm if this finding, which is based on a very limited amount of data, is widespread and valid.

The teachers in this study mainly had teaching experience in Arab countries, particularly Saudi Arabia. Thus, their attitudes may have been affected by working in this particular context. Further research could investigate teachers with experience in different contexts. Also, the assessment policies investigated in this study are used by a single university in Saudi Arabia. Other contexts may use different policies, and thus teachers coming from different contexts could be affected by different assessment policies. Accordingly, different contexts and the teachers coming from these contexts should be researched to find out if assessment polices are different from one context to another and if teachers are affected in a similar way or differently by these policies.

Finally, further research into EFL teachers' attitudes to intelligibility and nativeness should consider rebuilding the assessment of students' speech task and the acceptability task used in this study before adopting them as this study did not find both tasks strongly and positively correlated (see Appendix T), and thus this may cast doubt on the usefulness of such tasks. Further research may also need to make sure that the study's participants understand the terminology used in tasks

by, for example, providing previous training or education to the participants. The participants in the current study have not been provided with specific training or education on pronunciation issues or terminology as part of this study in advance of data collection. Further research should also consider shortening the assessment of students' speech task as much as possible since it took the participants almost more than an hour to complete based on the feedback of the pilot study participants. Shorting this task could help make it less cognitively demanding.

The recommendations mentioned in this section are some examples of the points that should be addressed in further research in the future. More and more aspects of the research into EFL teachers' attitudes towards intelligibility and nativeness in pronunciation assessment should be explored to obtain a deep understanding of this area of research and, as a result, provide recommendations, training and education accordingly.

References

Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning: A Journal of Research in Language Studies, 59*(2), 249-306. doi:10.1111/j.1467-9922.2009.00507.x

ACTFL. (2012). *ACTFL proficiency guidelines*. Alexandra, VA: Author. Retrieved July 7, 2020, from https://www.actfl.org/sites/default/files/guidelines/ACTFLProficiencyGuidelines2012.pdf

Agathopoulou, E. (2010). EFL student teachers' beliefs and the effect of a second language acquisition course. *Advances in Research on Language Acquisition and Teaching: Selected Papers*, 165-177. Retrieved April 14, 2020, from https://www.researchgate.net/publication/267704406_EFL_student_teachers'_beliefs_and_the_effect_of_a_second_language_acquisition_course/citations

Agheyisi, R., & Fishman, J. A. (1970). Language attitude studies: A brief survey of methodological approaches. *Anthropological Linguistics, 12*(5), 137-157.

Ahn, K. (2011). Conceptualization of American English native speaker norms: A case study of an English language classroom in south Korea. *Asia Pacific Education Review, 12*(4), 691-702. doi:10.1007/s12564-011-9169-6

Ajzen, I. & Fishbein, M. (1980). Understanding attitudes and predicting social behavior. Englewood Cliffs, NJ: Prentice Hall.

Ajzen, I. & Fishbein, M. (2005). The influence of attitudes on behavior. In D. Albarracı´n, B. T. Johnson, & M. P. Zanna (Eds.), Handbook of attitudes and attitude change (pp. 173-221). Hillsdale, NJ: Erlbaum.

Ajzen, I., & Timko, C. (1986). Correspondence between health attitudes and behavior. *Basic and Applied Social Psychology, 7*(4), 259-276. doi:10.1207/s15324834basp0704_2

Albaladejo, S. A. (2018). *The pronunciation of English as a lingua franca*. Manuscript in preparation.

Albarracín, D., Johnson, B. T., Fishbein, M., & Muellerleile, P. A. (2001). Theories of reasoned action and planned behavior as models of condom use: A meta-analysis. *Psychological Bulletin, 127*(1), 142-161. doi:10.1037/0033-2909.127.1.142

Al-Khouri, A. (2012). Population growth and government modernisation efforts: The case of GCC countries. *International Journal of Research in Management & Technology* (IJRMT), 2(1), 1-8. Retrieved August 18, 2020, from https://pdfs.semanticscholar.org/8593/431f3b7f1ca3d37f1999292a4af8212419ec.pdf

Alptekin, C. (2002). Towards intercultural communicative competence in ELT. *ELT Journal, 56*(1), 57-64. doi:10.1093/elt/56.1.57

Anderson, J. (1987). The markedness differentiational hypothesis and syllable structure difficulty. In G. Ioup & S. Winberger (Eds.), *Interlanguage phonology: The acquisition of second language sound system* (pp. 279-291). Cambridge, MA: Newbury House.

Andersonhsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning, 42*(4), 529-555. doi.org/10.1111/j.1467-1770.1992.tb01043.x

Angelovska, T. and Hahn, A. (2009). English with a native-like accent: an empirical study on proficiency. In A. G. Benati (Ed.). *Issues in second language proficiency* (pp. 147-166). Retrieved December 18, 2019, from https://ebookcentral.proquest.com

Arrieta, M. (2016). *Teacher and student perceptions of world Englishes (WE) pronunciations in two US settings.* MA Dissertation, Portland State University, OR. Retrieved January 12, 2018, from https://search-proquest-com.idpproxy.reading.ac.uk/docview/1880556763?pq-origsite=summon&accountid=13460

Arts - Rabigh & College of Sciences - English Language Department. (2016, March 23). *kau.edu.sa*. Retrieved December 17, 2020, from https://csar-eng.kau.edu.sa/Pages-about-dept.aspx

English Language Institute (2017, May 30). *kau.edu.sa*. Retrieved December 21, 2020, from https://eli.kau.edu.sa/Pages-Overview-er.aspx

Avery, P., & Ehrlich, S. (1992). *Teaching American English pronunciation*. Oxford: Oxford University Press.

Bachman, L. F., & Savignon, S. J. (1986). The evaluation of communicative language proficiency: A critique of the ACTFL oral interview. *The Modern Language Journal (Boulder, Colo.), 70*(4), 380-390. https://doi.org/10.1111/j.1540-4781.1986.tb05294.x

Baker, A. (2011). Discourse prosody and teachers' stated beliefs and practices. *TESOL Journal, 2*(3), 263-292. doi:10.5054/tj.2011.259955

Baker, A. (2013). Integrating fluent pronunciation use into content-based ESL instruction: Two case studies. In J. Levis, & K. LeVelle (Eds.), *Proceedings of the 4th pronunciation in second language learning and teaching conference* (pp. 245-254). Ames, IA: Iowa State University.

Baker, A. (2014). Exploring teachers' knowledge of second language pronunciation techniques: Teacher cognitions, observed classroom practices, and student perceptions. *TESOL Quarterly, 48*(1), 136-163. doi:10.1002/tesq.99

Baker, A. A. (2011a). Discourse prosody and teachers' stated beliefs and practices. *TESOL Journal, 2*(3), 263-292. doi:10.5054/tj.2011.259955

Baker, A. A. (2011b). ESL teachers and pronunciation pedagogy: Exploring the development of teachers' cognitions and classroom practices. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference, Sept. 2010.* (pp. 82-94). Aimes, IA: Iowa State University. Retrieved April 9, 2020, from https://www.academia.edu/12031228/Proccedings_of_the_2nd_Pronunciation_in_Second_Language_Learning_and_Teaching_Conference

Baker, A., & Burri, M. (2016). Feedback on second language pronunciation: A case study of EAP teachers' beliefs and practices. *Australian Journal of Teacher Education (Online), 41*(6), 1-20. doi:10.14221/ajte.2016v41n6.1

Baker, A., & Murphy, J. (2011). Knowledge base of pronunciation teaching: Staking out the territory. *TESL Canada Journal, 28*(2), 29. doi:10.18806/tesl.v28i2.1071

Baker, C. (1992). *Attitudes and language*. Clevedon, England: Multilingual Matters.

Ball, P. (1983). Stereotypes of Anglo-Saxon and non-Anglo-Saxon accents: Some exploratory Australian studies with the matched guise technique. *Language Sciences, 5*(2), 163-183. doi:10.1016/S0388-0001(83)80021-7

Ballard, L., & Winke, P. (2017). Students' attitudes towards English teachers' accents: The interplay of accent familiarity, comprehensibility, intelligibility, perceived native speaker status, and acceptability as a teacher. In T. Isaacs & P. Trofimovich (Eds.), *Second language pronunciation assessment: Interdisciplinary perspectives* (pp. 121-140). Bristol, England: Multilingual Matters. doi:10.21832/ISAACS6848

Bansal, R. K. (1990). The pronunciation of English in India. In S. Ramsaran (Ed.), *Studies in the pronunciation of English: A commemorative volume in honour of A. C. Gimson* (pp. 219-230). London, England: Routledge.

Barnwell, D. (1989). "Naive" native speakers and judgements of oral proficiency in Spanish. *Language Testing*, *6*(2), 152–163. https://doi.org/10.1177/026553228900600203

Barona, D.B. (2008). Native and non-native speakers' perceptions of non-native accents. Language and Literature Journal, 3(2). Retrieved June 29, 2021, from https://lljournal.commons.gc.cuny.edu/2008-2-vargas-barona-texto/

Basturkmen, H. (2012). Review of research into the correspondence between language teachers' stated beliefs and practices. *System, 40*(2), 282-295. doi:10.1016/j.system.2012.05.001

Basturkmen, H., Loewen, S., & Ellis, R. (2004). Teachers' stated beliefs about incidental focus on form and their classroom practices. *Applied Linguistics, 25*(2), 243-272. doi:10.1093/applin/25.2.243

Bayard, D., Weatherall, A., Gallois, C., & Pittam, J. (2001). Pax Americana? accent attitudinal evaluations in New Zealand, Australia and America. *Journal of Sociolinguistics, 5*(1), 22-49. doi:10.1111/1467-9481.00136

Beare, K. (2015*). How many people learn English globally?* Retrieved August 18, 2010, from http://esl.about.com/od/englishlearningresources/f/f_eslmarket.htm.

Bent, T., & Bradlow, R. (2003). The Interlanguage Speech Intelligibility Benefit. *Journal of the Acoustic Society of America*, *114*(3), 1600-1610. Retrieved November 22, 2017, from http://tessabent.com/interlanguage.pdf

Berger, C., & Bradac, J. J. (1982). Language and Social Knowledge: Uncertainty in Interpersonal Relations. London: Arnold.

Berns, M. (2008). World Englishes, English as a lingua franca, and intelligibility. *World Englishes, 27*(3-4), 327-334. doi:10.1111/j.1467-971X.2008.00571.x

Bernstein, J. (1999). *PhonePass testing: Structure and construct*. Menlo Park, CA: Ordinate Corporation. Retrieved June 18, 2020, from http://www.ordinate.com/pdf/StructureAndConstruct990826.pdf

Bernstein, J., Van Moere, A., & Cheng, J. (2010). Validating automated speaking tests. *Language Testing, 27*(3), 355-377. doi:10.1177/0265532210364404

Biricik Deniz, E., Kemaloglu-Er, E., & Ozkan, Y. (2020). ELF-aware pre-service teacher education: Practices and perspectives. ELT Journal, 74(4), 453-462. https://doi.org/10.1093/elt/ccaa040

Bley-Vroman, R. (1983). The comparative fallacy in interlanguage studies: The case of systematicity. *Language Learning, 33*(1), 1-17. doi:10.1111/j.1467-1770.1983.tb00983.x

Bøhn, H., & Hansen, T. (2017). Assessing pronunciation in an EFL context: Teachers' orientation towards nativeness and intelligibility. *Language Assessment Quarterly, 14*(1), 54-68. doi:10.1080/15434303.2016.1256407

Bohner, G., Wanke, M., & Michaela, W. (2002). *Attitudes and attitude change*. Retrieved April 11, 2020, from https://ebookcentral.proquest.com

Bokhorst-Heng, W. D., Alsagoff, L., Mckay, S., & Rubdy, R. (2007). English language ownership among Singaporean Malays: Going beyond the NS/NNS dichotomy. *World Englishes, 26*(4), 424-445. https://doi.org/10.1111/j.1467-971X.2007.00521.x

Bolton, K. (2004). World Englishes. In A. Davies & C. Elder (Eds.), *The handbook of applied linguistics* (pp. 369-396). Oxford, England: Blackwell.

Boyle, R. (2011). Patterns of change in English as a lingua franca in the UAE. *International Journal of Applied Linguistics, 21*(2), 143-161. doi:10.1111/j.1473-4192.2010.00262.x

Borg, S. (1999). Teachers' theories in grammar teaching. *ELT Journal, 53*(3), 157-167. doi:10.1093/elt/53.3.157

Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching, 36*(2), 81-109. doi:10.1017/S0261444803001903

Borg, S. (2006). *Teacher cognition and language education: Research and practice*. Retrieved April 9, 2020, from https://www.researchgate.net/publication/268438742_Teacher_Cognition_and_Language_Teacher_Education_beliefs_and_practice_A_conversation_with_Simon_Borg

Borg, S. (2011). The impact of in-service teacher education on language teachers' beliefs. *System, 39*(3), 370-380. doi:10.1016/j.system.2011.07.009

Boyd, S. (2003). Foreign-born teachers in the multilingual classroom in Sweden: The role of attitudes to foreign accent. *International Journal of Bilingual Education and Bilingualism, 6*(3-4), 283-295. doi:10.1080/13670050308667786

Bradac, J. J. (1990). Language attitudes and impression formation. In H. Giles & W. P. Robinson (Eds.), *Handbook of language and social psychology* (pp. 387-412). Chichester, England: Wiley.

Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition, 106*(2), 707-729. https://doi.org/10.1016/j.cognition.2007.04.005

Brannan, D., & Bleistein, T. (2012). Novice ESOL teachers' perceptions of social support networks. *TESOL Quarterly, 46*(3), 519-541. doi:10.1002/tesq.40

Breitkreutz, J. A., Derwing, T. M., & Rossiter, M. J. (2001). Pronunciation teaching practices in Canada. *TESL Canada Journal*, *19*, 51-61. Retrieved September 11, 2019, from https://teslcanadajournal.ca/index.php/tesl/article/download/919/738/

Bresnahan, M. J., Ohashi, R., Nebashi, R., Liu, W. Y., & Morinaga Shearman, S. (2002). Attitudinal and affective response toward accented English. *Language and Communication, 22*(2), 171-185. doi:10.1016/S0271-5309(01)00025-8

Brown, A. (1995). The effect of rater variables in the development of an occupation-specific language performance test. Language Testing, 12(1), 1–15. https://doi.org/10.1177/026553229501200101

Brown, J. D. (2012). EIL Curriculum Development. In Alsagoff, L., Hu, Guangwei, Mckay, S. L., & Renandya, W. A. (Eds.). *Principles and practices for teaching English as an International language* (pp. 147-167). New York: Routledge.

Brown, K. (2002). Ideology and context: World Englishes and EFL teacher training. *World Englishes, 21*(3), 445-448. doi:10.1111/1467-971X.00264

Brown, A., Iwashita, N., & McNamara, T. (2005). An examination of rater orientation and test- taker performance on English-for-academic-purposes speaking tasks. Princeton, NJ: Educational Testing Service.

Browne, K., & Fulcher, G. (2017). Pronunciation and Intelligibility in Assessing Spoken Fluency. In Isaacs T. & Trofimovich P. (Eds.), *Second Language Pronunciation Assessment: Interdisciplinary Perspectives* (pp. 37-53). Bristol, England: Multilingual Matters. Retrieved Aug 19, 2018, from http://www.jstor.org/stable/10.21832/j.ctt1xp3wcc.7

Brumfit, C. J. (2001). Individual Freedom in Language Teaching: Helping Learners to Develop a Dialect of their Own. Oxford: Oxford University Press.

Buck, G. (1989). Written tests of pronunciation: Do they work? *ELT Journal, 43*(1), 50-56. doi:10.1093/elt/43.1.50

Buckingham, L. (2015). Shades of cosmopolitanism: EFL teachers' perspectives on English accents and pronunciation teaching in the gulf. *Journal of Multilingual and Multicultural Development, 36*(6), 638-653. doi:10.1080/01434632.2014.994638

Bundgaard-Nielsen, R. L., Best, C. T., Kroos, C., & Tyler, M. D. (2012). Second language learners' vocabulary expansion is associated with improved second language vowel intelligibility. *Applied Psycholinguistics, 33*(3), 643-664. doi:10.1017/S0142716411000518

Burgess, J., & Spencer, S. (2000). Phonology and pronunciation in integrated language teaching and teacher education. *System, 28*(2), 191-215. doi:10.1016/S0346-251X(00)00007-5

Burns, A. (2006). Integrating research and professional development on pronunciation teaching in a national adult ESL program. *TESL Reporter*, *39*, 34-41. Retrieved December 9, 2019, from https://www.researchgate.net/publication/282664301_Integrating_research_and_professional_development_on_pronunciation_teaching_in_a_national_adult_ESL_program

Burri, M. (2015). Student teachers' cognition about L2 pronunciation instruction: A case study. *Australian Journal of Teacher Education (Online), 40*(10), 66-87. doi:10.14221/ajte.2015v40n10.5

Buss, L. (2013). Pronunciation from the perspective of pre-service EFL teachers: An analysis of internship reports. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference* (pp. 255–264). Retrieved September 11, 2019, from https://www.academia.edu/12031368/Proceedings_of_the_4th_Pronunciation_in_Second_Language_Learning_and_Teaching_Conference?auto=download

Buss, L. (2016). Beliefs and practices of Brazilian EFL teachers regarding pronunciation. *Language Teaching Research, 20*(5), 619-637. doi:10.1177/1362168815574145

Bayram, A. B. (2018). Serious subjects: A test of the seriousness technique to increase participant motivation in political science experiments. *Research & Politics, 5*(2), 205316801876745. doi:10.1177/2053168018767453

Byram, M. (1997). Teaching and Assessing Intercultural Communicative Competence. Clevedon: Multilingual Matters.

Byram, M., Gribkova, B., & Starkey, H. (2002). Developing the intercultural dimension in language teaching: A practical introduction for teachers. The Council of Europe. Retrieved October 19, 2020, from https://rm.coe.int/16802fc1c3

Canagarajah, A. S. (2006a). Changing communicative needs, revised assessment objectives: Testing English as an international language. *Language Assessment Quarterly, 3*(3), 229-242. doi:10.1207/s15434311laq0303_1

Canagarajah, A. S. (2006b). Negotiating the local in English as a lingua franca. *Annual Review of Applied Linguistics, 26*, 197-218. doi:10.1017/S0267190506000109

Canagarajah, A. S. (Ed.). (2004). *Reclaiming the local in language policy and practice*. Retrieved from https://ebookcentral.proquest.com

Carey, M. D., Mannell, R. H., & Dunn, P. K. (2011). Does a rater's familiarity with a candidate's pronunciation affect the rating in oral proficiency interviews? *Language Testing, 28*(2), 201. doi:10.1177/0265532210393704

Cargile, A. C. (2002). Speaker evaluation measures of language attitudes: Evidence of information-processing effects. *Language Awareness, 11*(3), 178-191. 10.1080/09658410208667055

Cargile, A. C., & Bradac, J. J. (2001). Attitudes toward language: a review of speaker- evaluation research and a general process model. In W. B. Gudykunst (Ed.), *Communication yearbook 25* (pp. 347-382). Mahwah, New Jersey: Lawrence Erlbaum Associates. Retrieved May 5, 2020, from https://epdf.pub/communication-yearbook-25.html

Cargile, A. C., & Giles, H. (1997). Understanding language attitudes: Exploring listener affect and identity. *Language & Communication, 17*(3), 195-217. doi:10.1016/s0271-5309(97)00016-5

Cargile, A. C., Giles, H., Ryan, E. B., & Bradac, J. J. (1994). Language attitudes as a social process: A conceptual model and new directions. *Language and Communication, 14*(3), 211-236. doi:10.1016/0271-5309(94)90001-9

Catford, J. C. (1987). Phonetics and the teaching of pronunciation: A systematic description of the teaching of English phonology. In J. Morley (Ed.), *Current perspectives on pronunciation: Practices anchored in theory* (pp. 83–100). Washington, DC: Teachers of English to Speakers of Other Languages.

CEFR Level A2. (n.d.). Stgiles-international. Retrieved February 25, 2021, from https://www.stgiles-international.com/app/webroot/docs/Level-A2-Learner-Outcomes.pdf

Celce-Murcia, M., Brinton, D., & Goodwin, J. (1996). *Teaching pronunciation: a reference for teachers of English to speakers of other languages*. Cambridge: Cambridge University Press.

Chalhoub-Deville, M., & Wigglesworth, G. (2005). Rater judgment and English language speaking proficiency. *World Englishes, 24*(3), 383-391. doi:10.1111/j.0083-2919.2005.00419.x

Chan, J. Y. H. (2016). A multi-perspective investigation of attitudes towards English accents in Hong Kong: Implications for pronunciation teaching. *TESOL Quarterly, 50*(2), 285-313. doi:10.1002/tesq.218

Chun, C. W. (2008). Comments on "evaluation of the usefulness of the versant for English test: A response": The author responds. *Language Assessment Quarterly, 5*(2), 168-172. doi:10.1080/15434300801934751

Clarke, C.M. and Garrett, M.F. (2004) Rapid adaptation to foreign-accented English. *Journal of the Acoustical Society of America* 116 (6), 3647–3658. Retrieved July 1, 2021, from https://www.researchgate.net/publication/8072143_Rapid_adaptation_to_foreign-accented_English

Cogo, A., & Dewey, M. (2012). *Analysing English as a lingua franca : A corpus-driven investigation*. Retrieved May 9, 2020, from https://ebookcentral.proquest.com

Cohen, L., Manion, L., & Morrison, K. (2018). Action research. In L. Cohen, L. Manion, & K. Morrison (Ed.). *Research methods in education* (8th ed., pp. 440-456). Retrieved December 12, 2020, from https://ebookcentral.proquest.com

Cook, V. (1999). Going beyond the native speaker in language teaching. *TESOL Quarterly, 33*(2), 185-209. doi:10.2307/3587717

Cook, V. (2007). *The goals of ELT Reproducing native speakers or promoting multi-competence among second language users.* London, England: Springer

Cook, V., & Singleton, D. (2014). *Key topics in second language acquisition*. Multilingual Matters.

Cortazzi, M., & Jin, L. (1999). Cultural mirrors: Materials and methods in the EFL classroom. In E. Hinkel (Ed.), *Culture in second language teaching and learning* (pp. 196-219). Cambridge. Cambridge University Press.

Coskun, A. (2011). Future English teachers' attitudes towards EIL pronunciation. *Journal of English as an International Language*, *6*(2), 46-68. Retrieved October 21, 2020, from https://files.eric.ed.gov/fulltext/ED527146.pdf

Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge, UK: Cambridge University Press.

Council of Europe. (2022, February 05). *Global scale - Table 1 (CEFR 3.3): Common Reference levels. https://www.coe.int/en/web/common-european-framework-reference-languages/table-1-cefr-3.3-common-reference-levels-global-scale*

Couper, G. (2003). The value of an explicit pronunciation syllabus in ESOL teaching. *Prospect, 18*, 53-70. Retrieved December 9, 2019, from https://pdfs.semanticscholar.org/c2fd/77a0c484eefd0570bd6a603016d505dccbd5.pdf?_ga=2.76914838.2042339453.1575898842-827565570.1575898842

Couper, G. (2006). The short and long-term effects of pronunciation instruction. *Prospect, 21*(1), 46–66. Retrieved September 16, 2019, from http://www.ameprc.mq.edu.au/docs/prospect_journal/volume_21_no_1/21_1_3_Couper.pdf

Couper, G. (2011). What makes pronunciation teaching work? Testing for the effect of two variables: Socially constructed metalanguage and critical listening. *Language Awareness, 20*(3), 159-182. doi:10.1080/09658416.2011.570347

Couper, G. (2016). Teacher cognition of pronunciation teaching amongst English language teachers in Uruguay. Journal of Second Language Pronunciation, 2(1), 29–55. doi:10.1075/jslp.2.1.02cou (Not reached)

Couper, G. (2017). Teacher cognition of pronunciation teaching: Teachers' concerns and issues. *TESOL Quarterly, 51*(4), 820-843. doi:10.1002/tesq.354

Coupland, N., & Bishop, H. (2007). Ideologised values for British accents. *Journal of Sociolinguistics, 11*(1), 74-93. doi:10.1111/j.1467-9841.2007.00311.x

Crowther, D., Trofimovich, P., Isaacs, T., & Saito, K. (2015). Does a speaking task affect second language comprehensibility? *The Modern Language Journal, 99*(1), 80-95. doi:10.1111/modl.12185

Cruttenden, A. (2014). *Gimson's pronunciation of English*. Retrieved from https://ebookcentral.proquest.com

Crystal, D. (2001). The Future of Englishes. In A. Burns, & C. Coffin (Eds.), *Analysing English in a Global Context: A Reader* (pp. 234-265). London: Routledge.

Crystal, D. (1997). English as a global language. Cambridge: Cambridge University Press.

Crystal, D. (2003). *English as a global language*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511486999

Crystal, D. (2008). Into the twenty-first century. In L. Mugglestone (Ed.), *The Oxford History of English* (pp. 394-413). Oxford: Oxford University Press.

Crystal, D. (2012). *English as a global language* (2nd ed.). Retrieved from https://ebookcentral.proquest.com

Da-Silva, R. (1998). A small-scale investigation into the intelligibility of the pronunciation of Brazilian intermediate students. *Speak Out! Newsletter of the IATEFL Pronunciation Special Interest Group, 23*, 19-25. Retrieved January 6, 2018, from http://hancockmcdonald.com/sites/hancockmcdonald.com/files/file-downloads/Ricardo%20Sili%20Intelligibility%20of%20Brazilians.pdf

Dalton-Puffer, C., Kaltenbock, G., & Smit, U. (1995). Language attitudes of L2 learners to native and non-native varieties of English. *Vienna English Working Paper*, *4*(4), 76-86. Retrieved February 18, 2018, from http://www.academia.edu/837758/Language_attitudes_of_L2_learners_to_native_and_non-native_varieties_of_English

Dauer, R. (1993). *Accurate English: A complete course in pronunciation*. Englewood Cliffs, NJ: Prentice Hall Regents. Retrieved December 10, 2019, from https://www.academia.edu/39045891/Accurate_English_a_Complete_Course_in_Pronunciation_C?auto=download

Dauer, R. M. (2005). The lingua franca core: A new model for pronunciation instruction? *TESOL Quarterly, 39*(3), 543-550. doi:10.2307/3588494

Davidson, F. (1993). Testing English across cultures: Summary and comments. *World Englishes, 12*(1), 113-125. doi:10.1111/j.1467-971X.1993.tb00013.x

Davidson, F. (2006). World Englishes and test construction. In B, Kachru, Y. Kachru & C. Nelson (Eds.), *The Handbook of World Englishes* (pp. 709-717). doi:10.1002/9780470757598.ch39

Davies, A. (2003). *The native speaker: Myth and reality*. Channel View Publications.

Davies, A., Hamp-Lyons, L., & Kemp, C. (2003). Whose norms? International proficiency tests in English. *World Englishes, 22*(4), 571-584. doi:10.1111/j.1467-971X.2003.00324.x

Davis, T. (2017). *Crash test dummies: Surprising lessons from the book of judges*. Abingdon Press. Retrieved November 15, 2021, from https://ebookcentral.proquest.com/lib/reading/detail.action?docID=4799980

Dávila, A., Bohara, A. K., & Saenz, R. (1993). Accent penalties and the earnings of Mexican Americans. *Social Science Quarterly, 74*(4), 902-916. Retrieved August 1, 2020, from http://web.b.ebscohost.com.idpproxy.reading.ac.uk/ehost/detail/detail?vid=0&sid=440a3258-bd98-4805-8592-773b26fec7a8%40sessionmgr103&bdata=JkF1dGhUeXBlPWlwLHNoaWIsdWlkJnNpdGU9ZWhvc3QtbGl2ZQ%3d%3d#AN=9406223050&db=bth

Dewaele, J. M. (2018). Why the Dichotomy 'L1 Versus LX User' is Better than 'Native Versus Non-native Speaker.' *Applied Linguistics*, *39*(2), 236–240.

De Jong, N. H., & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods, 41*(2), 385-390. doi:10.3758/BRM.41.2.385

De Klerk, V., & Bosch, B. (1995). Linguistic stereotypes: Nice accent — nice person? *International Journal of the Sociology of Language, 116*(1), 17-38. doi:10.1515/ijsl.1995.116.17

Denzin, N. K. (1997). Triangulation in educational research. In P. K. John (Ed.), *Educational research, methodology and measurement: An international handbook* (pp. 318-323). Oxford, England: Pergamon.

Derwing, T. (2008) Curriculum issues in teaching pronunciation to second language learners. In J. Hansen Edwards & M. Zampini (Eds.), *Phonology and second language acquisition* (347-369). Retrieved December 31, 2019, from Retrieved from https://ebookcentral.proquest.com

Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition, 19*(1), 1-16. Retrieved November 9, 2017, from https://www-cambridge-org.idpproxy.reading.ac.uk/core/journals/studies-in-second-language-acquisition/article/accent-intelligibility-and-comprehensibility/729C15F62F9EC9A51A33EAB5C2D05ED0

Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly, 39*(3), 379-398. doi:10.2307/3588486

Derwing, T. M., & Munro, M. J. (2009a). Comprehensibility as a factor in listener interaction preferences: Implications for the workplace. *The Canadian Modern Language Review / La Revue Canadienne Des Langues Vivantes, 66*(2), 181-202. doi:10.3138/cmlr.66.2.181

Derwing, T. M., & Munro, M. J. (2009b). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching, 42*(4), 476-490. doi:10.1017/S026144480800551X

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

Derwing, T. M., Munro, M. J., & Thomson, R. I. (2008). A longitudinal study of ESL learners' fluency and comprehensibility development. *Applied Linguistics, 29*(3), 359-380. doi:10.1093/applin/amm041

Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning, 48*(3), 393-410. doi:10.1111/0023-8333.00047

Derwing, T. M., & Rossiter, M. J. (2002). ESL learners' perceptions of their pronunciation needs and strategies. *System, 30*(2), 155-166. doi:10.1016/S0346-251X(02)00012-X

Deterding, D. (2010). ELF-based pronunciation teaching in China. *Chinese Journal of Applied Linguistics*, 33(6), 3–15.

Deterding, D. (2013). *Misunderstandings in English as a lingua franca: An analysis of elf interactions in south-east Asia*. Retrieved November 1, 2020, from https://ebookcentral.proquest.com

Deterding, D., & Gardiner, I. A. (2018). New pronunciation en route to world Englishes. In O. Kang, R. I. Thomson & J. M. Murphy (Eds), *The Routledge Handbook of Contemporary English Pronunciation* (pp. 218-231). Abingdon, Oxon: Routledge.

Deterding, D., & Kirkpatrick, A. (2006). Emerging South-East Asian Englishes and intelligibility. *World Englishes, 25*(3-4), 391-409. doi:10.1111/j.1467-971X.2006.00478.x

Deterding, D., & Lewis, C. (2019). Pronunciation in English as lingua franca. *Second handbook of English language teaching* (pp. 785-798). Springer International Publishing. https://doi.org/10.1007/978-3-030-02899-2_41

Dewaele, J. (2018). Why the dichotomy 'L1 versus LX user' is better than 'Native versus non-native speaker'. *Applied Linguistics, 39*(2), 236-240. https://doi.org/10.1093/applin/amw055

Dimova, S. (2018). Pronunciation assessment in the context of World Englishes. In O. Kang & A. Ginther (Eds), *Assessment in second language pronunciation* (pp. 49-66). Abingdon, England: Routledge.

Dörnyei, Z. (2007). Research Methods in Applied Linguistics. Oxford, England: Oxford University Press.

Downey, R., Farhady, H., Present-Thomas, R., Suzuki, M., & Van Moere, A. (2008). Evaluation of the usefulness of the versant for English test: A response. *Language Assessment Quarterly, 5*(2), 160-167. doi:10.1080/15434300801934744

Dubois, B., & Burns, J. A. (1975). An analysis of the meaning of the question mark response category in attitude scales. Educational and Psychological Measurement, 35(4), 869–884. doi:10.1177/001316447503500414

Dyers, C., & Abongdia, J. (2010). An exploration of the relationship between language attitudes and ideologies in a study of francophone students of English in Cameroon. *Journal of Multilingual and Multicultural Development, 31*(2), 119-134. doi:10.1080/01434630903470837

Dziubalska-Kołaczyk, K. (2005). Proceedings from PTLC 2005: *Native or Non-Native? This is the Question: Which English to Teach in the Globalizing World?* London: Department of Phonetics and Linguistics, UCL. Retrieved January 1, 2018, from https://www.ucl.ac.uk/pals/study/cpd/cpd-courses/ptlc/proceedings_2005/ptlcp67

Eagly, A. H. & Chaiken, S. (1998). Attitude structure and function. In D.T. Gilbert, S.T. Fiske & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 1, pp. 269-322). New York: McGraw-Hill.

Ekong, P. (1982). on the use of an indigenous model for teaching English in Nigeria. *World Englishes, 1*(3), 87-92. doi:10.1111/j.1467-971X.1982.tb00474.x

El-Dash, L. G., & Busnardo, J. (2001). Brazilian attitudes toward English: Dimensions of status and solidarity. *International Journal of Applied Linguistics, 11*(1), 57-74. https://doi.org/10.1111/1473-4192.00004

Elder, C., & Davies, A. (2006). Assessing English as a lingua franca. *Annual Review of Applied Linguistics, 26*, 282-304. doi:10.1017/S0267190506000146

Elder, C. and Harding, L. (2008) Language testing and English as an international language: Constraints and contributions. *Australian Review of Applied Linguistics* 31 (3), 34.1–34.11. Retrieved May 31, 2021, from https://benjamins.com/catalog/getpdf?uuid=aad2b3f0-2d10-4381-92b3-f02d1033814f&href=%2Faral%2Faral.31.3%2Faral.31.3.07eld%2Faral.31.3.07eld.pdf

Elliott, A. R. (1997). On the teaching and acquisition of pronunciation within a communicative approach. *Hispania, 80*(1), 95-108. doi:10.2307/345983

ETS. (2009). *The official guide to the TOEFL test* (3rd ed.). New York, NY: McGraw-Hill. Giles, H. (1979). Ethnicity markers in speech. In H. Giles, & K. R. Scherer (Eds.), *Social markers in speech* (pp. 251–290). Cambridge, England: Cambridge University Press.

ETS. (2010). *TOEIC® user guide: Speaking & writing*. Princeton, NJ: Educational Testing Service. Retrieved July 7, 2020, from https://www.ets.org/s/toeic/pdf/toeic-sw-score-user-guide.pdf

ETS. (2014). *TOEFL iBT® Test: Integrated speaking rubrics*. New York, NY: McGraw-Hill. Retrieved July 7, 2020, from https://www.ets.org/s/toefl/pdf/toefl_speaking_rubrics.pdf

Faez, F., & Valeo, A. (2012). TESOL teacher education: Novice teachers' perceptions of their preparedness and efficacy in the classroom. *TESOL Quarterly, 46*(3), 450-471. doi:10.1002/tesq.37

Fayer, J. M., & Krasinski, E. (1987). Native and nonnative judgments of intelligibility and irritation. *Language Learning, 37*(3), 313-326. doi:10.1111/j.1467-1770.1987.tb00573.x

Fazio, R. H. (1986). "How do attitudes guide behaviors?" In R.M. Sorrention & E.T. Higgins (Eds), The handbook of motivation and cognition: foundations of social behavior (Vol.1, pp. 204-243). New York: Guilford Press.

Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. In M. P. Zanna (Ed.), Advances in experimental social psychology (Vol. 23, pp. 75–107). New York: Academic.

Fazio, R. H. (1995). Attitudes as object-evaluation associations: Determinants, consequences, and correlates of attitude accessibility. In R. E. Petty & J. A. Krosnick (Eds.), Attitude strength: Antecedents and consequences (pp. 247–282). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Fazio, R. H., & Towles-Schwen, T. (1999). The MODE model of attitude-behavior processes. In S. Chaiken & Y. Trope (Eds.), Dual process theories in social psychology (pp. 97-116). New York: Guilford.

Fazio, R. H., Powell, M. C., & Williams, C. J. (1989). The role of attitude accessibility in the attitude-to-behavior process. *Journal of Consumer Research, 16*(3), 280-288. doi:10.1086/209214

Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly, 39*(3), 399-424. doi:10.2307/3588487

O'Connor, H., Madge, C., Shaw, R., & Wellens, J. (2008). Internet-based interviewing. In N. Fielding, R. M. Lee, & G. Blank (Eds.), *The Sage handbook of online research methods* (pp. 271–289). London: Sage. Retrieved March 17, 2021, from https://ebookcentral.proquest.com

Figueras, N., North, B., Takala, S., Avermaet, P., & Verhelst, N. (2009, January). Relating language examinations to the common European framework of reference for languages: learning, teaching, assessment (CEFR): A manual. https://rm.coe.int/1680667a2d

Flege, J. E. (1984). The detect of French accent by American listeners. *Journal of the Acoustical Society of America, 76*(3), 692-707. Retrieved July 16, 2020, from http://jimflege.com/files/Flege_French_accent_JASA_1984.pdf

Flege, J. E. (2003). Assessing constraints on second-language segmental production and perception. In A. Meyer, and N. Schiller (Eds.). (2003). *Phonetics and phonology in language comprehension and production: Differences and similarities* (319-55). Retrieved from https://ebookcentral.proquest.com

Flege, J. E., Munro, M. J., & Mackay, I. R. A. (1995). Effects of age of second-language learning on the production of English consonants. *Speech Communication, 16*(1), 1-26. doi:10.1016/0167-6393(94)00044-B

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language, 41*(1), 78-104. doi:10.1006/jmla.1999.2638

Flowerdew, J. (1994). research of relevance to second language lecture comprehension: An overview. In J. Flowerdew (Ed.), *Academic listening* (p. 7-29). Cambridge. Cambridge University Press.

Foote, J. A., Holtby, A. K., & Derwing, T. M. (2012). Survey of the teaching of pronunciation in adult ESL programs in Canada, 2010. *TESL Canada Journal, 29*(1), 1-22. doi:10.18806/tesl.v29i1.1086

Foote, J. A., Trofimovich, P., Collins, L., & Urzúa, F. S. (2016). Pronunciation teaching practices in communicative second language classes. *The Language Learning Journal, 44*(2), 181-196. doi:10.1080/09571736.2013.784345

Fraser, H. (2000). *Coordinating Improvements in Pronunciation Teaching for Adult Learners of English as a Second Language*. Canberra: DETYA (ANTA Innovative Project). Retrieved August 12, 2019, from https://helenfraser.com.au/wp-content/uploads/ANTA-REPORT-FINAL.pdf

Fraser, H. (2001). *Teaching pronunciation: A handbook for teachers and trainers: Three frameworks for an integrated approach*. Sydney, Australia: TAFE NSW Access Division. Retrieved December 7, 2019, from https://helenfraser.com.au/wp-content/uploads/HF-Handbook.pdf

Friedrich, P. (2000). English in brazil: Functions and attitudes. *World Englishes, 19*(2), 215-223. doi:10.1111/1467-971X.00170

Garret, P. (2010). *Attitudes to language*. Cambridge, England: Cambridge University Press. doi:10.1017/CBO9780511844713

Gass, S., & Varonis, E. M. (1984). The effect of familiarity on the comprehensibility of non-native speech. *Language Learning, 34*(1), 65-87. Retrieved July 31, 2020, from https://deepblue.lib.umich.edu/bitstream/handle/2027.42/98153/j.1467-1770.1984.tb00996.x.pdf?sequence=1

Gatbonton, E., Trofimovich, P., & Magid, M. (2005). Learners' ethnic group affiliation and L2 pronunciation accuracy: A sociolinguistic investigation. *TESOL Quarterly, 39*(3), 489-511. doi:10.2307/3588491

Gilbert, J. (2010). Pronunciation an orphan: What can be done? *SpeakOut: The Newsletter of the IATEFL Pronunciation Special Interest Group*, *43*, 3-7. Retrieved August 9, 2019, from https://www.tesol.org/docs/default-source/new-resource-library/pronunciation-as-orphan-what-we-can-do-about-it-.pdf?sfvrsn=0

Giles, H. (1972). The effect of stimulus mildness-broadness in the evaluation of accents. *Language and Speech, 15*(3), 262-269. doi:10.1177/002383097201500304

Giles, H. (1979). Ethnicity markers in speech. In H. Giles & K. R. Scherer (Eds.), *Social markers in speech* (pp. 251-290). Cambridge, England: Cambridge University Press.

Giles, H., & Billings, A. (2004). Assessing language attitudes: Speaker evaluation studies. In A. Davies & C. Elder (Eds.), *The handbook of applied linguistics* (pp. 187-209). Oxford, England: Blackwell.

Giles, H., Williams, A., Mackie, D. M., & Rosselli, F. (1995). Reactions to Anglo- and Hispanic-American-accented speakers: Affect, identity, persuasion, and the English-only controversy. *Language and Communication, 15*(2), 107-120. doi:10.1016/0271-5309(94)00019-9

Gluszek, A., & Dovidio, J. F. (2010). The way they speak: A social psychological perspective on the stigma of nonnative accents in communication. *Personality and Social Psychology Review, 14*(2), 214-237. doi:10.1177/1088868309359288

Golombek, P. R. (1998). A study of language teachers' personal practical knowledge. *TESOL Quarterly, 32*(3), 447-464. doi:10.2307/3588117

Gooauych, R., Saito, K., & Lyster, R. (2016). Effects of recasts and prompts on L2 pronunciation development: Teaching English to Korean adult EFL learners. *System, 60*, 117.

Graddol, D. (1997). The future of English?: A guide to forecasting the popularity of the English language in the 21st century. London: British Council.

Graddol, D. (2006). English next: Why global English may mean the end of 'English as a Foreign Language'. London, UK: The British Council.

Graden, E. C. (1996). How language teachers' beliefs about reading instruction are mediated by their beliefs about students. *Foreign Language Annals, 29*(3), 387-395. doi:10.1111/j.1944-9720.1996.tb01250.x

Grant, L. (2014). Prologue to the myths: What teachers need to know. In L. Grant (Ed.), *Pronunciation myths: Applying second language research to classroom teaching* (pp. 1-33). Ann Arbor, MI: The University of Michigan Press.

Graves, K. (2000). *Designing language courses*. Boston, MA: Heinle and Heinle. Retrieved May 8, 2020, from https://epdf.pub/designing-language-courses-a-guide-for-teachers.html

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis, 11*(3), 255-274. doi:10.2307/1163620

Greenleaf, J. H. (1929). French pronunciation tests. *The Modern Language Journal, 13*(7), 534-537. doi:10.1111/j.1540-4781.1929.tb00003.x

Griego-Jones, T. (1994). Assessing students' perceptions of biliteracy in two-way bilingual classrooms. *Journal of Educational issues of Language Minority Students, 13*, 79-93. Retrieved March 15, 2020, from https://www.researchgate.net/publication/265004446_Assessing_students%27_perceptions_of_biliteracy_in_two-way_bilingual_classrooms

Gurzynski-Weiss, L. (2013). Instructor characteristics and classroom-based SLA of Spanish. In K. L. Geeslin (Ed.), *The handbook of Spanish second language acquisition* (pp. 530-46). Chichester, West Sussex: John Wiley & Sons Inc. Retrieved May 8, 2020, from https://ebookcentral.proquest.com

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

Hamilton, J., Lopes, M., McNamara, T., & Sheridan, E. (1993). Rating scales and native speaker performance on a communicatively oriented EAP test. *Language Testing*, 10(3), 337–353. https://doi.org/10.1177/026553229301000307

Harding, L. (2012a). Accent, listening assessment and the potential for a shared-L1 advantage: A DIF perspective. *Language Testing, 29*(2), 163-180. doi:10.1177/0265532211421161

Harding, L. (2012b). Pronunciation assessment. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 4708-4713). Hoboken, NJ: John Wiley. doi:10.1002/9781405198431.wbeal0966

Harding, L. (2017). What do raters need in a pronunciation scale?: The user's view. In T. Isaacs & P. Trofimovich (Eds.), *Second language pronunciation assessment: Interdisciplinary perspectives* (pp. 12-34). Bristol, England: Multilingual Matters. doi:10.21832/ISAACS6848

Harding, L. (2018). Validity in pronunciation assessment. In O. Kang & A. Ginther (Eds), *Assessment in second language pronunciation*, (pp. 30-48). Abingdon, England: Routledge.

Harumi, I. (2002). A new framework of culture teaching for teaching English as a global language. *RELC Journal, 33*(2), 36-57. doi:10.1177/003368820203300202

Hayes-Harb, R., Smith, B. L., Bent, T., & Bradlow, A. R. (2008). The interlanguage speech intelligibility benefit for native speakers of mandarin: Production and perception of English word-final voicing contrasts. *Journal of Phonetics, 36*(4), 664-679. doi:10.1016/j.wocn.2008.04.002

He, D., & Miller, L. (2011). English teacher preference: The case of China's non-English-major students. *World Englishes, 30*(3), 428-443. 10.1111/j.1467-971X.2011.01716.x

Heather, D., & Carey, A. (2009). Teacher Beliefs. *The Gale Group.* Retrieved April 16, 2020, from https://www.education.com/articles/

Henderson, A., Frost, D., Tergujeff, E., Kautzsch, A., Murphy, D., Kirkova-Naskova, … Curnick, L. (2012). The English pronunciation teaching in Europe survey: Selected results. *Research in Language*, *10*(1), 1–23. Retrieved December 9, 2019, from https://www.researchgate.net/publication/270262770_The_English_pronunciation_teaching_in_Europe_survey_Selected_results

Hilgendorf, S. K. (2007). English in Germany: Contact, spread and attitudes. *World Englishes, 26*(2), 131-148. doi:10.1111/j.1467-971X.2007.00498.x

Holmes, J. (2008). An introduction to sociolinguistics (3rd ed.). Harlow, England: Pearson Longman.

Housen, A., & Kuiken, F. (2009). Complexity, accuracy and fluency in second language acquisition. *Applied Linguistics, 30*(4), 461-473. doi:10.1093/applin/amp048

Hrubes, D., Ajzen, I., & Daigle, J. (2001). Predicting hunting intentions and behavior: An application of the theory of planned behavior. *Leisure Sciences, 23*(3), 165-178. doi:10.1080/014904001316896855

Hsu, H. (2012). *The impact of world Englishes on language assessment: Rater attitude, rating behavior, and challenges* (Doctoral dissertation). Retrieved December 30, 2019, from https://search-proquest-com.idpproxy.reading.ac.uk/docview/1426157432?accountid=13460

Huang, B., Alegre, A., & Eisenberg, A. (2016). A cross-linguistic investigation of the effect of raters' accent familiarity on speaking assessment. *Language Assessment Quarterly, 13*(1), 25-41. doi:10.1080/15434303.2015.1134540

IELTS. (n.d.) *IELTS Speaking band descriptors* (public version). Retrieved July 7, 2020, from https://www.ielts.org/-/media/pdfs/speaking-band-descriptors.ashx?la=en

IELTS. (2008). IELTS speaking test: Instructions to IELTS examiners. Cambridge, England: IELTS.

Isaacs, T. (2008). Towards defining a valid assessment criterion of pronunciation proficiency in non-native English-speaking graduate students. *Canadian Modern Language Review, 64*(4), 555-580. doi:10.3138/cmlr.64.4.555

Isaacs, T. (2009). Integrating form and meaning in L2 pronunciation instruction. *TESL Canada Journal, 27*(1), 1. doi:10.18806/tesl.v27i1.1034

Isaacs, T. (2013). Assessing pronunciation. In A. J. Kunnan (Ed.), *The companion to language assessment* (pp. 140-155). Hoboken, NJ: Wiley-Blackwell. doi:10.1002/9781118411360.wbcla012

Isaacs, T. (2018a). Shifting sands in second language pronunciation teaching and assessment research and practice. *Language Assessment Quarterly: Conceptualizing and Operationalizing Speaking Assessment for a New Century, 15*(3), 273-293. doi:10.1080/15434303.2018.1472264

Isaacs, T. (2018b). Fully automated speaking assessment: Changes to proficiency testing and the role of pronunciation. In O. Kang, R. I. Thomson, & J. Murphy (Eds.), *The Routledge handbook of English pronunciation* (pp. 570–584). Abingdon, Oxon: Routledge.

Isaacs, T., & Harding, L. (2017). Pronunciation assessment. *Language Teaching, 50*(3), 347-366. doi:10.1017/S0261444817000118

Isaacs, T., & Thomson, R. I. (2013). Rater experience, rating scale length, and judgments of L2 pronunciation: Revisiting research conventions. *Language Assessment Quarterly, 10*(2), 135-159. doi:10.1080/15434303.2013.769545

Isaacs, T., & Trofimovich, P. (2011). Phonological memory, attention control, and musical ability: Effects of individual differences on rater judgments of second language speech. *Applied Psycholinguistics, 32*(1), 113-140. doi:10.1017/S0142716410000317

Isaacs, T., & Trofimovich, P. (2012). Deconstructing comprehensibility: Identifying the linguistic influences on listeners' L2 comprehensibility ratings. *Studies in Second Language Acquisition, 34*(3), 475-505. doi:10.1017/S0272263112000150

Isaacs, T., & Trofimovich, P. (2017). Key themes, constructs, and interdisciplinary perspectives in second language pronunciation assessment. In T. Isaacs & P. Trofimovich (Eds.), *Second language pronunciation assessment: Interdisciplinary perspectives* (pp. 3–11). Bristol, UK: Multilingual Matters. doi:10.21832/ISAACS6848

Isaacs, T., Trofimovich, P., & Foote, J. A. (2018). Developing a user-oriented second language comprehensibility scale for English-medium universities. *Language Testing, 35*(2), 193-216. doi:10.1177/0265532217703433

Jakobson, R. (1941). *Child language, aphasia and phonological universals*. The Hague, Netherlands: Mouton. Retrieved December 13, 2017, from https://ebookcentral.proquest.com/lib/reading/detail.action?docID=3044126

Jefferson, G. (1979). A technique for inviting laughter for its subsequent acceptance/declination. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology* (79-96). London, England: Wiley.

Jenkins, J. (1998). Which pronunciation norms and models for English as an international language? *ELT Journal, 52*(2), 119-126. doi:10.1093/elt/52.2.119

Jenkins, J. (2000). The phonology of English as an international language: New models, new norms, new goals. Oxford, England: Oxford University Press.

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

Jenkins, J. (2004). Research in teaching pronunciation and intonation. *Annual Review of Applied Linguistics, 24*, 109-125. doi:10.1017/S0267190504000054

Jenkins, J. (2005). Teaching pronunciation for English as a lingua franca: A sociopolitical perspective. In C. Gnutzmann & F. Intemann (Eds.), *The globalisation of English and the English language classroom.* (pp. 145-158). Göttingen, Germany: Gunter Narr. Retrieved January 6, 2018, from https://www.researchgate.net/publication/271642936_Teaching_pronunciation_for_English_as_a_Lingua_Franca_A_sociopolitical_perspective

Jenkins, J. (2006a). The spread of EIL: A testing time for testers. *ELT Journal, 60*(1), 42-50. doi:10.1093/elt/cci080

Jenkins, J. (2006b). Current perspectives on teaching world Englishes and English as a lingua franca. *TESOL Quarterly, 40*(1), 157-181. doi:10.2307/40264515

Jenkins, J. (2006c). Global Intelligibility and Local Diversity: Possibility or Paradox? In R. Rubdy & M. Saraceni (Eds.), *English in the World – Global Rules, Global Roles* (pp. 123-154). London: Continuum.

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

Jenkins, J. (2008). ELF at the gate: the position of English as a Lingua Franca. *Modern English Teacher*, *17*(2), 15. Retrieved May 30, 2021, from https://link.gale.com/apps/doc/A178412091/AONE?u=rdg&sid=summon&xid=254e95ff

Jenkins, J. (2009a). (Un)pleasant? (In)correct? (Un)intelligible? ELF speakers' perceptions of their accents. In A. Mauranen, & E. Ranta (Eds.), *English as a Lingua Franca: Studies and findings* (pp. 10-36). Newcastle, England: Cambridge Scholars Publishing. Retrieved January 1, 2018, from https://www.researchgate.net/publication/270283630_Unpleasant_Incorrect_Unintelligible_ELF_speakers%27_perceptions_of_their_accents

Jenkins, J. (2009b). Exploring attitudes towards English as a lingua franca in the East Asian context. In K. Murata & J. Jenkins (Eds.), *Global Englishes in Asian contexts: Current and future debates (*pp. 40-56)*.* Basingstoke, England: Palgrave Macmillan. Retrieved November 10, 2017, from https://www.researchgate.net/publication/270283569_Exploring_Attitudes_towards_English_as_a_Lingua_Franca_in_the_East_Asian_Context

Jenkins, J. (2009c). English as a lingua franca: Interpretations and attitudes. *World Englishes, 28*(2), 200-207. doi:10.1111/j.1467-971x.2009.01582.x

Jenkins, J. (2012). English as a lingua franca from the classroom to the classroom. *ELT Journal, 66* (4), 486-494. Retrieved September 11, 2019, from https://www.researchgate.net/publication/262243160_English_as_a_Lingua_Franca_from_the_classroom_to_the_classroom

Jenkins, J. (2014). *Global Englishes: A resource book for students*. Retrieved September 11, 2019, from https://ebookcentral.proquest.com

Jenkins, J., & Leung, C. (2013). English as a lingua franca. In A. J. Kunnan (Ed.), *The companion to language assessment* (pp. 1605-1616). Chichester, England: John Wiley & Sons. doi:10.1002/9781118411360.wbcla047

Jenkins, J., Cogo, A., & Dewey, M. (2011). Review of developments in research into English as a lingua franca. *Language Teaching, 44*(3), 281-315. doi:10.1017/S0261444811000115

Jindapitak, N. (2015). English as a lingua franca: Learners' views on pronunciation. *E-FLT: Electronic Journal of Foreign Language Teaching, 12*(2), 260. Retrieved February 5, 2018, from http://e-flt.nus.edu.sg/v12n22015/jindapitak.pdf

Jolanta S. K. (2018). Instructional models in the global context. In O. Kang, R. I. Thomson & J. M. Murphy (Eds), *The Routledge Handbook of Contemporary English Pronunciation* (pp. 232-246). Abingdon, Oxon: Routledge.

Kachru, B. B. (1985). Standards, codification and sociolinguistic realism: The English language in the outer circle. In R. Quirk, & H. G. Widdowson (Eds.), *English in the world: Teaching and learning the language and literatures* (pp. 11–30). Cambridge, England: Cambridge University Press.

Kachru, B. B. (1986). The alchemy of English: The spread functions and models of non-native Englishes. Urbana, IL: University of Illinois Press.

Kachru, B. B. (1992). *The other tongue: English across cultures* (2nd ed.). Urbana, IL: University of Illinois Press.

Kachru, B., & Nelson, C. (2001). World Englishes. In A. Burns & C. Coffin (Eds), *Analyzing English in a global context: A reader* (pp. 9-25). New York: Routledge.

Kachru, B. B., Kachru, Y., & Nelson, C. L. (Eds.). (2006). *The handbook of world Englishes*. Malden, MA: Blackwell.

Kachru, Y. (1994). Monolingual bias in SLA research. *TESOL Quarterly, 28*(4), 795-800. doi:10.2307/3587564

Kamiya, N. (2016). What effect does reading academic articles on oral corrective feedback have on ESL teachers? *TESOL Journal, 7*(2), 328-349. doi:10.1002/tesj.210

Kang, O. (2008). Ratings of L2 Oral Performance in English: Relative Impact of Rater Characteristics and Acoustic Measures of Accentedness. *Spaan Fellow Working Papers in Second or Foreign Language Assessment*, *6*, 181-205. Retrieved June 29, 2021, from https://www.researchgate.net/publication/251844787_Ratings_of_L2_Oral_Performance_in_English_Relative_Impact_of_Rater_Characteristics_and_Acoustic_Measures_of_Accentedness

Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System, 38*(2), 301-315. doi:10.1016/j.system.2010.01.005

Kang, O. (2012). Impact of rater characteristics and prosodic features of speaker accentedness on ratings of international teaching assistants' oral performance. *Language Assessment Quarterly, 9*(3), 249-269. doi:10.1080/15434303.2011.642631

Kang, O., & Ginther, A. (2018). Introduction. In O. Kang & A. Ginther (Eds), *Assessment in second language pronunciation* (pp. 1-7). Abingdon, England: Routledge.

Kang, O., & Kermad, A. (2018). Assessment in L2 pronunciation. In O. Kang, R. Thomson, & J. Murphy (Eds.). The Routledge Handbook of Contemporary English Pronunciation (pp. 511-526). Abingdon, England: Routledge. Retrieved Aug 9, 2018, from https://www.scribd.com/document/374398927/The-Routledge-Handbook-of-Contemporary-English-Pronunciation

Kang, O., Rubin, D., & Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *The Modern Language Journal, 94*(4), 554-566. doi:10.1111/j.1540-4781.2010.01091.x

Kang, O., Thomson, R. I., & Moran, M. (2018). Empirical approaches to measuring the intelligibility of different varieties of English in predicting listener comprehension. *Language Learning, 68*(1), 115-146. doi:10.1111/lang.12270

Kaulfers, W. V. (1944). Wartime development in modern-language achievement testing. *The Modern Language Journal, 28*(2), 136-150. doi:10.1111/j.1540-4781.1944.tb04835.x

Kennedy, S. (2011). L2 proficiency: measuring the intelligibility of Words and Extended Speech. In A. G. Benati (Ed.). *Issues in second language proficiency* (pp. 132-146). Retrieved December 18, 2019, from https://ebookcentral.proquest.com

Kennedy, S., & Trofimovich, P. (2008). Intelligibility, comprehensibility, and accentedness of L2 speech: The role of listener experience and semantic context. *Canadian Modern Language Review, 64*(3), 459-489. doi:10.3138/cmlr.64.3.459

Kennedy, S., Blanchet, J., & Guénette, D. (2017). Teacher-raters' assessment of French lingua franca pronunciation. In T. Isaacs & P. Trofimovich (Eds.), *Second language pronunciation assessment: Interdisciplinary perspectives* (pp. 210-236). Bristol, UK: Multilingual Matters. doi:10.21832/ISAACS6848

Kenworthy, J. (1987). *Teaching English pronunciation*. London: Longman.

Kerckvoorde, C. v. (2001). *Caught between norms: The English pronunciation of Dutch learners.* Linguistic Society of America. doi:10.1353/lan.2001.0089

Khaghaninejad, M. S., & Maleki, A. (2015). The effect of explicit pronunciation instruction on listening comprehension: Evidence from Iranian English learners. *Theory and Practice in Language Studies, 5*(6), 1249. doi:10.17507/tpls.0506.18

Kharma, N., & Hajjaj, A. (1997). Errors in English Among Arabic Speakers: Analysis and Remedy. Beirut: York Press.

Khatib, M., & Monfared, A. (2017). Exploring English teachers' attitudes towards pronunciation issues and varieties of English in three circles of world Englishes. *Applied Research on English Language, 6*(2), 213-236. doi:10.22108/are.2017.21349

Kim, Y. (2009). An investigation into native and non-native teachers' judgments of oral English performance: A mixed methods approach. *Language Testing, 26*(2), 187-217. doi:10.1177/0265532208101010

Kim, H., & Billington, R. (2018). Pronunciation and comprehension in English as a lingua franca communication: Effect of L1 influence in international aviation communication. *Applied Linguistics, 39*(2), 135-158. doi:10.1093/applin/amv075

Kirkpatrick, A. (2006). Which Model of English: Native-speaker, Nativized or Lingua Franca? In R. Rubdy & M. Saraceni (Eds.), *English in the World – Global Rules, Global Roles* (pp. 423-534). London: Continuum.

Kirkpatrick, A. (2007). World Englishes: Implications for international communication and English language teaching. Cambridge: Cambridge University Press.

Kirkpatrick, A. (2010a). *The Routledge handbook of world Englishes*. London, England: Routledge.

Kirkpatrick, A. (2010b). *English as a lingua franca in ASEAN. A Multilingual model*. Hong Kong, China: Hong Kong University Press. Retrieved October 24, 2017, from https://ebookcentral.proquest.com

Kirkpatrick, A. (2011). English as an asian lingua franca and the multilingual model of ELT. *Language Teaching, 44*(2), 212-224. doi:10.1017/S0261444810000145

Kramsch, C. (1993). *Context and culture in language teaching*. Oxford: Oxford University Press.

Krashen, S. (1981). Second language acquisition and second language learning. Oxford, UK: Pergamon Press.

Kraut, R., & Wulff, S. (2013). Foreign-accented speech perception ratings: A multifactorial case study. *Journal of Multilingual and Multicultural Development, 34*(3), 249-263. doi:10.1080/01434632.2013.767340

Krezanoski, P. J., Bangsberg, D. R., & Tsai, A. C. (2018). Quantifying bias in measuring insecticide-treated bednet use: meta-analysis of self-reported vs objectively measured adherence. *Journal of Global Health*, *8*(1), 010411–010411

Kubanyiova, M., & Feryok, A. (2015). Language teacher cognition in applied linguistics research: Revisiting the territory, redrawing the boundaries, reclaiming the relevance. *The Modern Language Journal, 99*(3), 435-449. doi:10.1111/modl.12239

Kumaravadivelu, B. (2012). *Language teacher education for a global society*. Routledge: Taylor& Francis.

Kvale, S. (2007). Qualitative research kit: Doing interviews. London: Sage.

Labov, W. (1964). *The Social Stratification Of English In New York City* (Doctoral dissertation). Retrieved from May 31, 2021, from https://www.proquest.com/dissertations-theses/social-stratification-english-new-york-city/docview/302149177/se-2?accountid=13460 (302149177)

Ladegaard, H. J., & Sachdev, I. (2006). 'I like the Americans. but I certainly don't aim for an American accent': Language attitudes, vitality and foreign language learning in Denmark. *Journal of Multilingual and Multicultural Development, 27*(2), 91-108. doi:10.1080/01434630608668542

Lado, R. (1961). Language testing: The construction and use of foreign language tests. London, UK: Longman.

Lambert, W. E. (1967). A social psychology of bilingualism. *Journal of Social Issues, 23*(2), 91-109. doi:10.1111/j.1540-4560.1967.tb00578.x

Lee, H. (2012). World Englishes in a High School English Class: A Case from Japan. In Matsuda. A. (Ed.). *Principles and practices for teaching English as an International language* (pp. 154-168). Bristol: Multilingual Matters.

Lee, I. (2009). Ten mismatches between teachers' beliefs and written feedback practice. *ELT Journal, 63*(1), 13-22. doi:10.1093/elt/ccn010

Lee, J., Jang, J., & Plonsky, L. (2015). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics, 36*(3), 345-366. doi:10.1093/applin/amu040

Lee, W., & Ridley, A. (1999). What implications does English globalization have for treatment of students' spoken errors? *RELC Journal, 30*(2), 18-37. doi:10.1177/003368829903000202

Lenski, G. E., & Leggett, J. C. (1960). Caste, class, and deference in the research interview. *The American Journal of Sociology, 65*(5), 463-467. https://doi.org/10.1086/222750

Levis, J. M. (1999). Intonation in theory and practice, revisited. *TESOL Quarterly, 33*(1), 37-63. doi:10.2307/3588190

Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly, 39*(3), 369-378. doi:10.2307/3588485

Libben, G., & Rossman-Benjamin, T. (1992). TESL methodology in Canada: A study of instructor attitudes and correlates. *TESL Canada Journal*, *9*(2), 9-29. Retrieved January 10, 2021, from https://pdfs.semanticscholar.org/26a9/f87a31ed06447f55a020e42eea7a238b64a2.pdf

Litman, D., Strik, H., & Lim, G. S. (2018). Speech technologies and the assessment of second language speaking: Approaches, challenges, and opportunities. *Language Assessment Quarterly: Conceptualizing and Operationalizing Speaking Assessment for a New Century, 15*(3), 294-309. doi:10.1080/15434303.2018.1472265

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry.* Beverly Hills, CA: Sage.

Lindemann, S. (2002). Listening with an attitude: A model of native-speaker comprehension of non-native speakers in the United States. *Language in Society, 31*(3), 419-441. https://doi.org/10.1017/S0047404502020286

Lindemann, S. (2003). Koreans, Chinese or Indians? attitudes and ideologies about non-native English speakers in the united states. *Journal of Sociolinguistics, 7*(3), 348-364. doi:10.1111/1467-9481.00228

Lindemann, S. (2017). Variation or 'error'?: Perception of pronunciation variation and implications for assessment. In T. Isaacs & P. Trofimovich (Eds.), *Second language pronunciation assessment: Interdisciplinary perspectives* (pp. 193-209). Bristol, England: Multilingual Matters. doi:10.21832/ISAACS6848

Lippi-Green, R. (2012). English with an accent: Language, ideology, and discrimination in the United States (2nd ed.). Abingdon, England: Routledge.

Litzenberg, J. (2013). *An investigation of pre-service English language teacher attitudes towards varieties of English in interaction* (Doctoral dissertation). Retrieved February 14, 2018, from ProQuest Dissertations & Theses Global. (1466023541).

Liu, M., & Zhang, L. (2007). Student perceptions of native and non-native English teachers' attitudes, teaching skills assessment and performance. *Asian EFL Journal*, 9(4), 157-166. Retrieved from May 10, 2020, from https://www.asian-efl-journal.com/main-journals/student-perceptions-of-native-non-native-english-teachers-attitudes-teaching-skills-assessment-and-performance/

Liu, X. (2010). Arousing the College Students' Motivation in Speaking English Through Role-Play. *International Education Studies, 3*(1), 136. Retrieved April 16, 2020, from https://www.semanticscholar.org/paper/Arousing-the-College-Students-'-Motivation-in-Liu/5af8c9db6e9554f2265f3582fcf564c4cbd365dd

Lo, B. P. J., Orton, J., & Yihong, P. G. (2009). *China and English: Globalisation and the dilemmas of identity*. Retrieved May 10, 2020, from https://ebookcentral.proquest.com

Loewen, S. (2014). *Introduction to instructed second language acquisition*. Retrieved December 16, 2019, from https://ebookcentral.proquest.com

Lowenberg, P. H. (1993). Issues of validity in tests of English as a world language: Whose standards? *World Englishes, 12*(1), 95-106. doi:10.1111/j.1467-971X.1993.tb00011.x

Lowenberg, P. H. (2002). Assessing English proficiency in the expanding circle. *World Englishes, 21*(3), 431-435. doi:10.1111/1467-971X.00261

Lumley, T., & McNamara, T. F. (1995). Rater characteristics and rater bias: implications for training. Language Testing, 12(1), 54–71. https://doi.org/10.1177/026553229501200104

Lundeberg, O. K. (1929). Recent developments in audition-speech tests. *The Modern Language Journal, 14*(3), 193-202. doi:10.1111/j.1540-4781.1929.tb01268.x

Luoma, S. (2004). *Assessing speaking*. Cambridge, UK: Cambridge University Press.

Macdonald, S. (2002). Pronunciation views and practices of reluctant teachers. *Prospect*, *17*(3), 3-18. Retrieved August 12, 2019, from https://www.researchgate.net/publication/285517759_Pronunciation-views_and_practices_of_reluctant_teachers

Mackey, A., Mackey, A., Gass, S. M., & Gass, S. M. (2005). *Second language research: Methodology and design*. Retrieved February 1, 2021, from https://ebookcentral.proquest.com

Mahmoud, A. (2000). Modern standard Arabic vs. non-standard Arabic: Where do Arab students of EFL transfer from? *Language, Culture and Curriculum, 13*(2), 126-136. doi:10.1080/07908310008666594

Major, R. C., Fitzmaurice, S. F., Bunta, F., & Balasubramanian, C. (2002). The effects of nonnative accents on listening comprehension: Implications for ESL assessment. *TESOL Quarterly, 36*(2), 173-190. doi:10.2307/3588329

Matsuda, A. (2003a). The ownership of English in Japanese secondary schools. *World Englishes, 22*(4), 483-496. doi:10.1111/j.1467-971X.2003.00314.x

Matsuda, A. (2003b). Incorporating world englishes in teaching english as an international language. *TESOL Quarterly, 37*(4), 719-729. doi:10.2307/3588220

Matsuda, A. (2012). Teaching Material in EIL. In Alsagoff, L., Hu, Guangwei, Mckay, S. L., & Renandya, W. A. (Eds.). *Principles and practices for teaching English as an International language* (pp.168-185). New York: Routledge.

Matsuda, A., & Friedrich, P. (2011). English as an international language: A curriculum blueprint. *World Englishes, 30*(3), 332-344. doi:10.1111/j.1467-971X.2011.01717.x

Matsumoto, Y. (2011). Successful ELF communications and implications for ELT: Sequential analysis of ELF pronunciation negotiation strategies. The Modern Language Journal (Boulder, Colo.), 95(1), 97-114. https://doi.org/10.1111/j.1540-4781.2011.01172.x

Matsuura, H., Chiba, R., & Fujieda, M. (1999). Intelligibility and comprehensibility of American and Irish Englishes in Japan. World Englishes, 18(1), 49-62. https://doi.org/10.1111/1467-971X.00121

McArthur, T. (1987). The English languages? *English Today, 3*(3), 9-13. doi:10.1017/S0266078400013511

McCrocklin, S., & Link, S. (2016). Accent, identity, and a fear of loss? ESL students' perspectives. *Canadian Modern Language Review, 72*(1), 122-148. 10.3138/cmlr.2582

McDermott, R. (2011). Internal and External Validity. In J. N. Druckman, D. P. Green, J. H. Kuklinski, & A. Lupia (Eds). *Cambridge Handbook of Experimental Political Science* (pp. 27-40). New York: Cambridge University Press. doi:10.1017/CBO9780511921452

McKay, S. L. (2002). *Teaching English as an international language*. Oxford: Oxford University Press.

McKay, S. L. (2003). EIL curriculum development. *RELC Journal, 34*(1), 31-47. doi:10.1177/003368820303400103

McKay, S. L. (2006). EIL Curriculum Development. In R. Rubdy & M. Saraceni (Eds.), *English in the World – Global Rules, Global Roles* (pp. 34-65). London: Continuum.

McKay, S. L. (2012). Principles of teaching English as an international language. In L. Alsagoff, S. L. McKay, G. Hu, & W. A. Renandya (Eds.), *Principles and practices for teaching English as an international language* (pp. 28-46). New York: Routledge.

McKay, S., & Bokhorst-Heng, W. (2008). International English in its sociolinguistic contexts: Towards a socially sensitive EIL pedagogy. New York, NY: Routledge.

McKenzie, R. M. (2008). Social factors and non-native attitudes towards varieties of spoken English: A Japanese case study. *International Journal of Applied Linguistics, 18*(1), 63-88. doi:10.1111/j.1473-4192.2008.00179.x

McNamara, T. (1996). Measuring second language performance. London: Longman.

McNamara, T. (2012). Language assessments as shibboleths: A poststructuralist perspective. *Applied Linguistics, 33*(5), 564-581. doi:10.1093/applin/ams052

McNamara, T., & Roever, C. (2006). *Language Testing: The Social Dimension*. Oxford: Blackwell.

Mertens, D. M. (2015). Research and evaluation in education and psychology: integrating diversity with quantitative, qualitative and mixed methods (4th ed.). London: Sage.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). Qualitative data analysis: A methods sourcebook (3rd ed.). Thousand Oaks, CA: Sage Publications.

Miller, G. A. (1994). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 101*(2), 343-352. doi:10.1037/0033-295X.101.2.343

Milroy, J., (2007). The ideology of standard language, in C. Llamas, L. Mullany and P. Stockwell (eds.), *The Routledge companion to sociolinguistics*, (pp. 133–9). London: Routledge.

Miyagi, K. (2006). Japanese EFL teachers' perceptions of nonnative varieties of English: Are they ready to include other Englishes in their classrooms? (Publication No. 304931899) [Master's dissertation, McGill University). Available from ProQuest One Academic. Retrieved May 4, 2021, from https://search.proquest.com/dissertations-theses/japanese-efl-teachers-perceptions-nonnative/docview/304931899/se-2?accountid=13460

Modiano, M. (2009). Inclusive/exclusive? English as a lingua franca in the European Union. World Englishes, *28*(2), 208-223. Retrieved July 23, 2020, from https://www.researchgate.net/publication/230326984_Inclusiveexclusive_English_as_a_lingua_franca_in_the_European_Union

Monfared, A., & Khatib, M. (2018). English or Englishes? Outer and expanding circle teachers' awareness of and attitudes towards their own variants of English in ESL/EFL teaching contexts. *Australian Journal of Teacher Education (Online), 43*(2), 56-75. doi:10.14221/ajte.2018v43n2.4

Monfared, A., & Safarzadeh, M. M. (2014). Pronunciation issues and varieties of English from an EIL perspective: A survey of outer and expanding circle learners' beliefs. *International Journal of Applied Linguistics & English Literature, 3*(6), 212-223. 10.7575/aiac.ijalel.v.3n.6p.212

Mora, J. C., & Darcy, I. (2017). The relationship between cognitive control and pronunciation in a second language. In T. Isaacs & P. Trofimovich (Eds.), *Second language pronunciation assessment: Interdisciplinary perspectives* (pp. 95–120). Bristol, UK: Multilingual Matters. doi:10.21832/ISAACS6848

Morin, R. (2007). A neglected aspect of the standards: Preparing foreign language Spanish teachers to teach pronunciation. *Foreign Language Annals, 40*(2), 342-360. doi:10.1111/j.1944-9720.2007.tb03206.x

Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly, 25*(3), 481-507. doi:10.2307/3586981

Morley, J. (ed.) (1994). Pronunciation pedagogy and theory: new views, new directions. Alexandria, Virginia: TESOL.

Morris-Adams, M. (2016). Negotiating topic changes: native and non-native speakers of English in conversation. *International Journal of Applied Linguistics*, *26*(3), 366–383.

Moser, S. C., Kalton, G. (1993). *Survey methods in social investigation* (2nd ed.). Aldershot: Dartmouth.

Moussu, L., & Llurda, E. (2008). Non-native English-speaking English language teachers: History and research. *Language Teaching, 41*(3), 315-348. doi:10.1017/S0261444808005028

Moyer, A. (1999). Ultimate attainment in L2 phonology. *Studies in Second Language Acquisition, 21*(1), 81-108. doi:10.1017/S0272263199001035

Moyer, A. (2013). *Foreign accent: The phenomenon of non-native speech*. Cambridge, UK: Cambridge University Press. doi:10.1017/CBO9780511794407

Munro, M. J., & Derwing, T. M. (1995a). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 45*(1), 73-97. doi:10.1111/j.1467-1770.1995.tb00963.x

Munro, M. J., & Derwing, T. M. (1995b). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech, 38*(3), 289. Retrieved July 13, 2020, from https://search.proquest.com/docview/1299110416?accountid=13460

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. doi:10.1111/0023-8333.49.s1.8

Munro, M. J., & Derwing, T. M. (2011). The foundations of accent and intelligibility in pronunciation research. *Language Teaching, 44*(3), 316-327. doi:10.1017/S0261444811000103

Munro, M. J., & Derwing, T. M. (2015). Intelligibility research and practice: Teaching priorities. In M. Reed & J. M. Levis (Eds.), *The handbook of English pronunciation* (pp. 375-396). Malden, MA: Wiley Blackwell. doi:10.1002/9781118346952.ch21

Munro, M. J., Derwing, T. M., & Burgess, C. S. (2010). Detection of nonnative speaker status from content-masked speech. *Speech Communication, 52*(7), 626-637. doi:10.1016/j.specom.2010.02.013

Munro, M. J., Derwing, T. M., & Morton, S. L. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition, 28*(1), 111-131. doi:10.1017/S0272263106060049

Munro, M. J. (2003). A primer on accent discrimination in the Canadian context. *TESL Canada Journal, 20*, 38-51. Retrieved December 9, 2019, from https://files.eric.ed.gov/fulltext/EJ669737.pdf

Munro, M. J. (2018). Dimensions of pronunciation. In O. Kang, R. I. Thomson, & J. Murphy (Eds.), *The Routledge handbook of English pronunciation* (pp. 570–584). Abingdon, Oxon: Routledge.

Murphy, D. (2011). An investigation of English pronunciation teaching in Ireland. *English Today, 27*(4), 10-18. doi:10.1017/S0266078411000484

Murphy, J. M. (1997). Phonology courses offered by MATESOL programs in the U.S. *TESOL Quarterly, 31*(4), 741-764. doi:10.2307/3587758

Murphy, J. M. (2014b). Myth 7: Teacher training programs provide adequate preparation in how to teach pronunciation. In L. Grant (Ed.), *Pronunciation myths: Applying second language research to classroom teaching* (pp. 188-224). Ann Arbor, MI: The University of Michigan Press.

Murray, H. (2003). Swiss English teachers Euro-English: attitudes to a non-native variety. Bulletin suisse de linguistique appliquée, 77, 147-165. Retreived July 2, 2021, from https://core.ac.uk/download/pdf/20650501.pdf

Nagle, C., Sachs, R., & Zárate–Sández, G. (2018). Exploring the intersection between teachers' beliefs and research findings in pronunciation instruction. *The Modern Language Journal, 102*(3), 512-532. doi:10.1111/modl.12493

Nargis, SS. M. H. (2018). A study of teachers' belief about second language acquisition. *The Asian EFL Journal, 20*(1), 57-64. Retrieved April 14, 2020, from https://www.asian-efl-journal.com/monthly-journals/2018-teaching-articles/volume-20-issue-1-2018/

Nelson, C. L. (1982). Intelligibility and nonnative varieties of English. In B. B. Kachru (Ed.), *The other tongue: English across cultures* (pp. 58-73). Urbana, IL: University of Illinois Press.

Nelson, G. L. (1998). Intercultural communication and related courses taught in TESOL masters' degree programs. *International Journal of Intercultural Relations, 22*(1), 17-33. doi:10.1016/S0147-1767(97)00032-1

Nelson, C. L. (2008). Intelligibility since 1969. *World Englishes, 27*(3-4), 297-308. doi:10.1111/j.1467-971X.2008.00568.x

Nelson, C. L. (2012). *Intelligibility in world Englishes: Theory and application*. Florence: Routledge Ltd. doi:10.4324/9780203832578

Newbold, D. (2017). "Co-certification": A close encounter with ELF for an international examining board. Journal of English as a Lingua Franca, 6(2), 367-388. https://doi.org/10.1515/jelf-2017-0017

Newman, D. (2002). The phonetic status of Arabic within the world's languages: The uniqueness of the lughat al-daad. *Antwerp Papers in Linguistics*, 100, 65-75. Retrieved August 19, 2020, from http://dro.dur.ac.uk/4420/

North, B. (2000). The development of a common framework scale of language proficiency. New York: Peter Lang.

Nowacka, M. (2012). Questionnaire-based pronunciation studies: Italian, Spanish and Polish students' views on their English pronunciation. *Research in Language, 10*(1), 43-61. 10.2478/v10015-011-0048-3

Nymeyer, K., Dewey, D. P., Eggington, W., & Baker-Smemoe, W. (2022). Factors that affect native English speakers' comfort levels when communicating with non-native English speakers. *International Journal of Applied Linguistics*, *32*(1), 158–174.

O'Brien, M. G. (2014). L2 learners' assessments of accentedness, fluency, and comprehensibility of native and nonnative German speech. *Language Learning, 64*(4), 715-748. doi:10.1111/lang.12082

Ockey, G. J., & French, R. (2016). From one to multiple accents on a test of L2 listening comprehension. *Applied Linguistics, 37*(5), 693-715. doi:10.1093/applin/amu060

Oppenheim, A. N. (1992). Questionnaire design, interviewing and attitude measurement. London, England: Pinter.

Oral English Proficiency Program. (n.d.). *OEPT2 holistic scale*. Retrieved July 28, 2020, from https://www.purdue.edu/oepp/documents/OEPT%20HOLISTIC%20SCALE%2011-1-2016.pdf

Osimk, R. (2009). Decoding sounds: An experimental approach to intelligibility in ELF. *Vienna English Working Papers, 18*(1), 64-89. Retrieved November 5, 2017, from http://anglistik.univie.ac.at/fileadmin/user_upload/dep_anglist/weitere_Uploads/Views/0901final.pdf

Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using SPSS*. Retrieved February 3, 2021, from https://ebookcentral.proquest.com

Pantos, A. J., & Perkins, A. W. (2013). Measuring Implicit and Explicit Attitudes Toward Foreign Accented Speech. *Journal of Language and Social Psychology*, *32*(1), 3–20.

Pennington, M. C., & Ellis, N. C. (2000). Cantonese speakers' memory for English sentences with prosodic cues. *The Modern Language Journal, 84*(3), 372-389. doi:10.1111/0026-7902.00075

Phan Le Ha (2008). Teaching English as an international language: Identity, resistance and negotiation. Clevedon, UK: Multilingual Matters.

Phipps, A. (2013). Intercultural ethics: Questions of methods in language and intercultural communication. *Language and Intercultural Communication, 13*(1), 10-26. https://doi.org/10.1080/14708477.2012.748787

Phipps, S., & Borg, S. (2009). Exploring tensions between teachers' grammar teaching beliefs and practices. *System, 37*(3), 380-390. doi:10.1016/j.system.2009.03.002

Phothongsunan, S., & Suwanarak, K. (2008). Native and non-native dichotomy: Distinctive stances of Thai teachers of English. *Abac Journal*, *28*(2), 10-30. Retrieved March 18, 2018, from www.assumptionjournal.au.edu/index.php/abacjournal/article/download/540/486

Pickering, L. (2001). The role of tone choice in improving ITA communication in the classroom. *TESOL Quarterly, 35*(2), 233-255. doi:10.2307/3587647

Pickering, L. (2009). Intonation as a pragmatic resource in ELF interaction. Intercultural Pragmatics, 6(2), 235-255. Retrieved December 16, 2017, from http://faculty.tamuc.edu/lpickering/Pdfs/Publish_2.pdf

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? interpreting effect sizes in L2 research. Language Learning, 64(4), 878-912. https://doi.org/10.1111/lang.12079

Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica, 104*(1), 1-15. doi:10.1016/S0001-6918(99)00050-5

Prodromou, L. (2006). Defining the 'successful bilingual speaker' of English. In R. Rubdy & M. Saraceni (eds.), *English in the world* (pp. 51-70). London: Continuum.

Punch, K. F. (2014). Introduction to social research: Quantitative and qualitative approaches. London: Sage.

Purpura, J. E. (2016). Second and foreign language assessment. *The Modern Language Journal, 100*(1), 190-208. doi:10.1111/modl.12308

Rajadurai, J. (2007). Intelligibility studies: A consideration of empirical and ideological issues. *World Englishes, 26*(1), 87-98. doi:10.1111/j.1467-971X.2007.00490.x

Ramineni, C., Trapani, C. S., Williamson, D. M., Davey, T., & Bridgeman, B. (2012). evaluation of the e-rater® scoring engine for the TOFEL® independent and integrated prompts. *ETS Research Report Series, 2012*(1), i-51. doi:10.1002/j.2333-8504.2012.tb02288.x

Randall, M., & Samimi, M. A. (2010). The status of English in Dubai. *English Today, 26*(1), 43-50. doi:10.1017/S0266078409990617

Rang, O., & Moran, M. (2014). Functional loads of pronunciation features in nonnative speakers' oral assessment. *TESOL Quarterly, 48*(1), 176-187. doi:10.1002/tesq.152

Reeves, J. R. (2006). Secondary teacher attitudes toward including English-language learners in mainstream classrooms. *The Journal of Educational Research, 99*(3), 131-143. doi:10.3200/JOER.99.3.131-143

Richards, J. C. (1996). Teachers' maxims in language teaching. *TESOL Quarterly, 30*(2), 281-296. doi:10.2307/3588144

Richards, J. C., & Lockhart, C. (2000). *Reflective teaching in second language classrooms*. Cambridge: Cambridge University Press.

Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching* (3rd ed.). Cambridge, UK: Cambridge University Press.

Richards, J. C., Tung, P., & Ng, P. (1992). The Culture of the English Language Teacher: A Hong Kong Example. RELC Journal, 23(1), 81–102. https://doi.org/10.1177/003368829202300106

Rindal, U. (2010). Constructing identity with L2: Pronunciation and attitudes among Norwegian learners of English. *Journal of Sociolinguistics, 14*(2), 240-261. doi:10.1111/j.1467-9841.2010.00442.x

Riney, T. J., Takada, M., & Ota, M. (2000). Segmentals and global foreign accent: The Japanese flap in EFL. *TESOL Quarterly, 34*(4), 711-737. doi:10.2307/3587782

Robinson, P. (1996). Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions. *Studies in Second Language Acquisition, 18*(1), 27-67. doi:10.1017/S0272263100014674

Rolland, L., Dewaele, J. M., & Costa, B. (2020). Planning and conducting ethical interviews: power, language and emotions. In J. McKinley, & H. Rose (Eds.). *The Routledge Handbook of Research Methods in Applied Linguistics*. Routledge. https://doi.org/10.4324/9780367824471

Rubdi, R., & Saraceni, M. (2006). *English in the world: Global rules, global roles*. London: Continuum.

Rubin, D. L. (1992). Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. *Research in Higher Education, 33*(4), 511-531. doi:10.1007/BF00973770

Rubin, D. L., Healy, P., Gardiner, T. C., Zath, R. C., & Moore, C. P. (1997). Nonnative physicians as message sources: Effects of accent and ethnicity on patients' responses to AIDS prevention counseling. *Health Communication, 9*(4), 351-368. doi:10.1207/s15327027hc0904_4

Ryan, E. B., Giles, H., and Sebastian, R. J. (1982). An integrative perspective for the study of attitudes toward language variation. In H. Giles & E. Bouchard Ryan (Eds.), *Attitudes towards language variation: Social and applied contexts* (pp. 1-19). London, UK: Edward Arnold.

Saito, K. (2012). Effects of instruction on L2 pronunciation development: A synthesis of 15 quasi-experimental intervention studies. *TESOL Quarterly, 46*(4), 842-854. doi:10.1002/tesq.67

Saito, K. (2014). Experienced teachers' perspectives on priorities for improved intelligible pronunciation: The case of Japanese learners of English. *International Journal of Applied Linguistics, 24*(2), 250-277. doi:10.1111/ijal.12026

Saito, K., & Lyster, R. (2012). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /?/ by Japanese learners of English. *Language Learning: A Journal of Research in Language Studies, 62*(2), 595. https://doi.org/10.1111/j.1467-9922.2011.00639.x

Saito, K., Trofimovich, P., & Isaacs, T. (2015). Using listener judgments to investigate linguistic influences on L2 comprehensibility and accentedness: A validation and generalization study. *Applied Linguistics,* , amv047. doi:10.1093/applin/amv047

Saito, K., Trofimovich, P., & Isaacs, T. (2016). Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. *Applied Psycholinguistics, 37*(2), 217-240. doi:10.1017/S0142716414000502

Saito, K., Trofimovich, P., Isaacs, T., & Webb, S. (2017). Re-examining Phonological and Lexical Correlates of Second Language Comprehensibility: The Role of Rater Experience. In T. Isaacs & P. Trofimovich (Eds.), *Second language pronunciation assessment: Interdisciplinary perspectives* (pp. 141–155). Bristol, UK: Multilingual Matters. doi:10.21832/ISAACS6848

Sakai, S., & D'Angelo, F. J. (2005). A vision for world Englishes in the expanding circle. *World Englishes, 24*(3), 323-327. doi:10.1111/j.0883-2919.2005.00414.x

Sandelowski, M. (2003). Tables of rableaux? The challenges of writing and reading mixed methods studies. In A. Tashakkori & C. Reddlie (Eds.), *Handbook of mixed methods in social and behavioral research*. Thousands Oaks, CA: Sage.

Sannes, M. T. (2013). *From the Native Speaker Norm towards English as an International Language* (Master's dissertation). Retrieved December 29, 2019, from http://bora.uib.no/bitstream/handle/1956/6669/106843915.pdf?sequence=1

Sannes, M. T. (2013). *From the Native Speaker Norm towards English as an International Language* (Master's dissertation). Retrieved December 29, 2019, from http://ora.uib.no/bitstream/handle/1956/6669/106843915.pdf?sequence=1

Sato, K. (1998). *Evaluative reactions towards "foreign accented" English speech: The effects of listeners' experience on their judgements* (Order No. MQ28987). Available from ProQuest One Academic. (304467732). Retrieved August 1, 2020, from https://search.proquest.com/docview/304467732?accountid=13460

Scheuer, S. 2005. "Why Native Speakers Are (Still) Relevant? In K. Dziubalska-Kolaczyk & J. Przedlacka (Eds.), *English pronunciation models. A changing scene* (pp. 111-130). Bern, Switzerland: Peter Lang.

Schaetzel, K., & Low, E. L. (2009). Teaching pronunciation to adult English language learners. *CAELA Network Briefs*. Retrieved December 7, 2019, from http://www.cal.org/adultesl/pdfs/teaching-pronunciation-to-adult-english-ll.pdf

Schiffrin, D. (1987). *Discourse markers*. Cambridge, England: Cambridge University Press.

Schiffrin, D. (1994). *Approaches to discourse*. Oxford, England: Blackwell.

Schneider, E. W. (2007). *Postcolonial English: varieties around the world*. doi:10.1017/CBO9780511618901

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. Harvard Educational Review, 57(1), 1-21. Retrieved January 10, 2021, from https://people.ucsc.edu/~ktellez/shulman.pdf

Schmid, M. S. (2013). First language attrition. *Wiley Interdisciplinary Reviews. Cognitive Science*, *4*(2), 117–123.

Scott, R., & Rodgers, B. (1995). Changing teachers' conceptions of teaching writing: A collaborative study. *Foreign Language Annals, 28*(2), 234-246. doi:10.1111/j.1944-9720.1995.tb00788.x

Scovel, T. (1969). Foreign accents, language acquisition, and cerebral dominance 1.*Language Learning, 19*(3-4), 245-253. doi:10.1111/j.1467-1770.1969.tb00466.x

Seidlhofer, B. (2001). Closing a conceptual gap: The case for a description of English as a Lingua Franca. *International Journal of Applied Linguistics*, 11(2), 133–158.

Seidlhofer, B. (2004). Research perspectives on teaching English as a lingua franca. *Annual Review of Applied Linguistics, 24*, 209-239. doi:10.1017/S0267190504000145

Seidlhofer, B. (2006). English as a Lingua Franca in the Expanding Circle: What it Isn't. In R. Rubdy & M. Saraceni (Eds.), *English in the World – Global Rules, Global Roles*. London: Continuum.

Seidlhofer, B. (2009). Common ground and different realities: World Englishes and English as a lingua franca. *World Englishes, 28*(2), 236-245. doi:10.1111/j.1467-971X.2009.01592.x

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

Setter, J., & Jenkins, J. (2005). State-of-the-art review article. *Language Teaching, 38*(1), 1-17. doi:10.1017/S026144480500251X

Sewell, A. (2017). Pronunciation assessment in Asia's World City: Implications of a lingua franca approach in Hong Kong. In T. Isaacs & P. Trofimovich (Eds.), *Second language pronunciation assessment: Interdisciplinary perspectives* (pp. 237-255). Bristol, England: Multilingual Matters. doi:10.21832/ISAACS6848

Sharifian, D. F. (Ed.). (2009). *English as an international language: Perspectives and pedagogical issues*. Retrieved May 9, 2020, from https://ebookcentral.proquest.com

Sifakis, N. C., & Sougari, A. (2005). Pronunciation issues and EIL pedagogy in the periphery: A survey of Greek state school teachers' beliefs. *TESOL Quarterly, 39*(3), 467-488. doi:10.2307/3588490

Sifakis, N. (2007). The education of teachers of English as a lingua franca: A transformative perspective. *International Journal of Applied Linguistics, 17*(3), 355-375. doi:10.1111/j.1473-4192.2007.00174.x

Singleton, D. (2005). The critical period hypothesis: A coat of many colours. *International Review of Applied Linguistics in Language Teaching, 43*(4), 269-285. doi:10.1515/iral.2005.43.4.269

Smith, L. E. (1983). *Readings in English as an international language*. Oxford, England: Pergamon.

Smith, L. E. (1992). Spread of English and issues of intelligibility. In B. B. Kachru (Ed.), *The other tongue: English across cultures* (2nd ed.) (pp. 75-90). Urbana, IL: University of Illinois Press.

Smith, L. E. (2015). English as an international auxiliary language. *Journal of English as a Lingua Franca, 4*(1), 159-164. doi:10.1515/jelf-2015-0001

Smith, L. E., & Bisazza, J. A. (1982). the comprehensibility of three varieties of English for college students in seven countries. *Language Learning, 32*(2), 259-269. doi:10.1111/j.1467-1770.1982.tb00971.x

Smith, L. E., & Nelson, C. L. (1985). International intelligibility of English: Directions and resources. *World Englishes, 4*(3), 333-342. doi:10.1111/j.1467-971X.1985.tb00423.x

Smith, L. E., & Nelson, C. L. (2006). World Englishes and issues of intelligibility. In B. B. Kachru, Y. Kachru, & C. L. Nelson (Eds.), *The handbook of World Englishes* (pp. 428-445). Malden, MAs: Blackwell.

Smith, L. E., & Rafiqzad, K. (1979). English for cross-cultural communication: The question of intelligibility. *TESOL Quarterly, 13*(3), 371-380. doi:10.2307/3585884

Snow, M. A., Kamhi-Stein, L. D., & Brinton, D. M. (2006). Teacher training for English as a lingua franca. *Annual Review of Applied Linguistics, 26*, 261-281. doi:10.1017/S0267190506000134

Sobkowiak, W. (2005). Why not LFC? In K. Dziubalska-Kolaczyk & J. Przedlacka (Eds.), *English pronunciation models. A changing scene* (pp. 131-150). Bern, Switzerland: Peter Lang.

Sowden, C. (2012). ELF on a mushroom: The overnight growth in English as a lingua franca. *ELT Journal, 66*(1), 89-96. doi:10.1093/elt/ccr024

Spolsky, B. (1993). Testing across cultures: An historical perspective. *World Englishes, 12*(1), 87-93. doi:10.1111/j.1467-971X.1993.tb00010.x

Spolsky, B. (1995). Measured words: The development of objective language testing. Oxford, UK: Oxford University Press.

Sridhar, S. N. (1994). A reality check for SLA theories. *TESOL Quarterly, 28*(4), 800-805. doi:10.2307/3587565

Stanojević, M., Borenić, V. K., & Smojver, V. J. (2012). Combining different types of data in studying attitudes to English as a lingua franca. *Research in Language, 10*(1), 29-41. doi:10.2478/v10015-011-0043-8

Steinbach, M., & Kazarloga, V. (2014). Square-headed frogs and world citizens: Attitudes and identities of ESL teacher candidates in Québec. *Journal of Language, Identity & Education, 13*(5), 319-334. doi:10.1080/15348458.2014.958038

Stemler, S. E. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research & Evaluation, 9*(4), 1-11. doi:10.7275/96jp-xz07

Swan, M., & Smith, B. (2001). *Learner English: A teacher's guide to interference and other feature*. Cambridge: Cambridge University Press.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston: Pearson Education.

Tan, P. K. W., Ooi, V. B. Y., & Chiang, A. K. L. (2006). World Englishes or English as a Lingua Franca? A view from the Perspective of Non-Anglo Englishes. In R. Rubdy & M. Saraceni (Eds.), *English in the World – Global Rules, Global Roles*. London: Continuum.

Tannen, D. (1989). Talking voices repetition dialogue and imagery in conversation discourse. Cambridge, England: Cambridge University Press.

Tarone, E. (1987). The phonology of interlanguage. In G. Ioup, & S. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 70-85). Cambridge, MA: Newbury House.

Taylor, L. (2006). The changing landscape of English: Implications for language assessment. *ELT Journal, 60*(1), 51-60. doi:10.1093/elt/cci081

Taylor, L. (Ed.). (2011). Examining speaking: Research and practice in assessing second language speaking. Cambridge, UK: UCLES/Cambridge University Press.

Testing for normality using spss statistics. (n.d.). Statisitcs.Laerd. Retrieved February 1, 2021, from https://statistics.laerd.com/spss-tutorials/testing-for-normality-using-spss-statistics.php

Thompson, I. (1991). Foreign accents revisited: The English pronunciation of Russian immigrants. *Language Learning, 41*(2), 177-204. doi:10.1111/j.1467-1770.1991.tb00683.x

Thomson, R. I. (2011). Computer assisted pronunciation training: Targeting second language vowel perception improves pronunciation. *CALICO Journal, 28*(3), 744-765. doi:10.11139/cj.28.3.744-765

Thomson, R. (2018). Measurement of pronunciation constructs. In O. Kang & A. Ginther (Eds), *Assessment in second language pronunciation* (pp. 11-29). Abingdon, England: Routledge.

Thomson, R. I., & Derwing, T. M. (2015). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics, 36*(3), 326-344. doi:10.1093/applin/amu076

Timmis, I. (2002). Native-speaker norms and international English: A classroom view. *ELT Journal, 56*(3), 240-249. https://doi.org/10.1093/elt/56.3.240

Timmis, I. (2018). Pronunciation future in the twenty-first-century English-speaking world: context, choice and consultation. In O. Kang, R. I. Thomson & J. M. Murphy (Eds), *The Routledge Handbook of Contemporary English Pronunciation* (pp. 495-510). Abingdon, Oxon: Routledge.

Trafimow, D., & Sheeran, P. (1998). Some tests of the distinction between cognitive and affective beliefs. *Journal of Experimental Social Psychology, 34*(4), 378-397. doi:10.1006/jesp.1998.1356

Trofimovich, P., & Isaacs, T. (2012). Disentangling accent from comprehensibility. *Bilingualism: Language and Cognition,15*(4), 905-916. doi:10.1017/S1366728912000168

Trofimovich, P., & Isaacs, T. (2017). Second language pronunciation assessment: A look at the present and the future. In T. Isaacs & P. Trofimovich (Eds.), *Second language pronunciation assessment: Interdisciplinary perspectives* (pp. 259 -271). Bristol, England: Multilingual Matters. doi:10.21832/ISAACS6848

Ufomata, T, (1990). Acceptable models for TEFL (with special reference to Nigeria). In S. Ramsaran (Ed.), *Studies in the pronunciation of English: A commemorative volume in honour of A.C. Gimson* (pp. 212-216). London, England: Routledge.

Van Moere, A., & Suzuki, M. (2018). Using speech processing technology in assessing pronunciation. In O. Kang & A. Ginther (Eds), *Assessment in second language pronunciation*, (pp. 137-152). Abingdon, England: Routledge.

Varonis, E. M., & Gass, S. (1982). The comprehensibility of non-native speech. *Studies in Second Language Acquisition, 4*(2), 114-136. doi:10.1017/S027226310000437X

Walker, R. (2001). Pronunciation priorities, the lingua franca core, and monolingual groups. *Speak Out! Newletter of the IATEFL Pronunciation Special Interest Group, 28*, 4-9. Retrieved February 5, 2022, from https://associates.iatefl.org/pages/materials/ltskills19.pdf

Walker, R. (2010). *Teaching the pronunciation of English as a lingua franca*. Oxford, England: Oxford University Press.

Walker, R., Low, E. L., & Setter, J. (2021). *English pronunciation for a global world*. Oxford, England: Oxford University Press.

Warren, P., Elgort, I., & Crabbe, D. (2009). Comprehensibility and prosody ratings for pronunciation software development. *Language Learning and Technology, 13*(3), 87-102. Retrieved September 16, 2019, from https://scholarspace.manoa.hawaii.edu/bitstream/10125/44193/1/13_03_warrenelgortcrabbe.pdf

Watanabe, Y. (2013). The National Center Test for University Admissions. Language Testing, 30, 565-573. Retrieved June 16, 2020, from https://www.researchgate.net/publication/275576407_The_National_Center_Test_for_University_Admissions

Watson, J. C. E. (2002). *The phonology and morphology of Arabic*. ProQuest Ebook Central https://ebookcentral.proquest.com

Weigle, S. C. (1998). Using FACETS to model rater training effects. *Language Testing, 15*(2), 263-287. https://doi.org/10.1177/026553229801500205

Weinberger, S. (2013). Speech Accent Archive. Retrieved from http://accent.- gmu.edu

Wennerstrom, A. (2001). The music of everyday speech: prosody and discourse analysis. Oxford: Oxford University Press.

Whitaker, T., Whitaker, B., & Lumpa, D. (2009). *Motivating & inspiring teachers: The educational leader's guide for building staff morale*. Abingdon, Oxon: Routledge. doi:10.4324/9781315856445

Widdowson, H. G. (1994). The ownership of English. *TESOL Quarterly, 28*(2), 377-389. doi:10.2307/3587438

Winke, P., Gass, S., & Myford, C. (2013). Raters' L2 background as a potential source of bias in rating oral performance. *Language Testing, 30*(2), 231-252. doi:10.1177/0265532212456968

Winters, S., & O'Brien, M. G. (2013). Perceived accentedness and intelligibility: The relative contributions of F0 and duration. *Speech Communication, 55*(3), 486-507. doi:10.1016/j.specom.2012.12.006

Woods, D. (1996). *Teacher Cognition in Language Teaching*. Cambridge: Cambridge University Press.

Xi, X. (2010). Automated scoring and feedback systems: Where are we and where are we heading? *Language Testing, 27*(3), 291-300. doi:10.1177/0265532210364643

Xi, X., & Mollaun, P. (2011). Using raters from India to score a Large-Scale speaking test. *Language Learning, 61*(4), 1222-1255. doi:10.1111/j.1467-9922.2011.00667.x

Yamaguchi, C. (2002). Towards international English in EFL classroom in Japan. *The Internet TESL Journal*, *8*(1). Retrieved January 2, 2018, from http://iteslj.org/Articles/Yamaguchi-Language.html

Yan, X. (2014). An examination of rater performance on a local oral English proficiency test: A mixed-methods approach. *Language Testing, 31*(4), 501-527. doi:10.1177/0265532214536171

Yan, X., & Ginther, A. (2018). Listeners and raters: similarities and differences in evaluation of accented speech. In O. Kang & A. Ginther (Eds), *Assessment in second language pronunciation* (pp. 67-88). Abingdon, England: Routledge.

Yano, Y. (2001). World Englishes in 2000 and beyond. *World Englishes, 20*(2), 119-132. doi:10.1111/1467-971X.00204

Yorkey, R. (1974). Practical EFL techniques for teaching Arabic-speaking students. *International Review of Applied Linguistics in Language Teaching*, 37(4), 291-320. Retrieved August 19, 2020, from https://pdfs.semanticscholar.org/f0dd/16f16b40702d4cca999614c3ed423153eb0b.pdf?_ga=2.186969640.140218231.1597819004-238563229.1597819004

Young, T. J., & Walsh, S. (2010). Which English? whose English? an investigation of 'non-native' teachers' beliefs about target varieties. *Language, Culture and Curriculum, 23*(2), 123-137. doi:10.1080/07908311003797627

Zhang, Y., & Elder, C. (2011). Judgments of oral proficiency by non-native and native English speaking teacher raters: Competing or complementary constructs? *Language Testing, 28*(1), 31-50. doi:10.1177/0265532209360671

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

Zoghbor, W. (2011). The effectiveness of the Lingua Franca Core (LFC) in improving the perceived intelligibility and perceived comprehensibility of Arab learners at post-secondary level (Doctoral dissertation). Retrieved December 29, 2017, from https://lra.le.ac.uk/bitstream/2381/9635/1/2011ZOGHBORWSEDD.pdf

Zoghbor, W. (2014). English varieties and Arab learners in the Gulf Cooperation Council (GCC) Countries: attitude and perception. *Arab World English Journal*, 5(2), 167-186. Retrieved August 18, 2020, from https://www.researchgate.net/publication/291970522_English_Varieties_and_Arab_Learners_in_the_Gulf_Cooperation_Council_GCC_Countries_Attitude_and_Perception

Zoghbor, W. S. (2018). Teaching English pronunciation to multi-dialect first language learners: The revival of the lingua franca core (LFC). *System, 78*, 1-14. doi:10.1016/j.system.2018.06.008

Appendices

Appendix A: Speaking Assessment Guidelines of KAU's Preparatory Year English Language Program

Speaking Assessment Guidelines

ELI
JCEA

General Notes

- The first assessor (section teacher) acts as an interlocutor and grader and the second assessor (the non-section teacher) acts as an observer and rater. There is no interaction between the second assessor and the student or the interlocutor.
- For Task One, if after asking any question the student seems to not understand the question, the interlocutor (section teacher) may briefly paraphrase the question taking care not to give any examples beyond those provided in the prompt. This should only be done **in English**.
- For Task Two, students may only ask questions to the interlocutor in the **2 minute** preparation time given **prior to** the commencement of the role play task. Once again, the interlocutor may paraphrase any part of the question the student seems to not understand taking care not to give any examples beyond those provided in the prompt. All questions from students must be **asked and answered in English only.**
- Both assessors are provided with a separate set of task sheets, a Speaking Exam- PER ASSESSOR Score Recording Sheet, rating scales and these guidelines.
- The first and second assessors record their scores on the Speaking Exam- PER ASSESSOR Score Recording Sheet without discussing them.

Conducting the Assessment

- The interlocutor calls 2 students in the order of the class list to conduct the assessment.
- The assessors write the students' names and academic numbers on their copy of the "Speaking Exam PER ASSESSOR Score Recording Sheet".
- The classroom teacher also makes sure that the students write their name, ID number, **and sign** the "Speaking Exam FINAL Score Recording and Attendance Sheet"
- The interlocutor conducts the assessment according to the task-specific instructions provided below.
- At the end of Task 1, using the assessment rating scale, both assessors record the grade for "Task Achievement (Task One)" **only** on the Speaking Exam- PER ASSESSOR Score Recording Sheet.
- At the end of Task 2, using the assessment rating scale, both assessors record the grade for "Task Achievement (Task Two)". For levels 101 and 102 they also record the grade for "Interaction (Task Two)". Then, they record the grade for the **other** rubric categories based on the student's performance in **both Task 1 and Task 2.**
- Once the assessment is over, the interlocutor calls the next 2 students and the same procedure is repeated.

Task 1: Short interview questions

- The assessors have been provided with 10 sets of 5 prompts (lettered A-J) on one sheet of A4 paper. The assessors should cycle through the prompts **in order** (the first student gets Set A and the second student gets Set B. Then, when the next pair is called, the first student in the new pair would get Set C and the second student would get Set D, etc).
- Although the assessors are calling up 2 students at a time, **each** student will be asked a **different** set of prompts in Task One. Depending on the size of the class, each set of prompts may be used 3-4 times in a cyclical fashion.
- The interlocutor should read aloud the short introduction to the questions and then start asking the student each of the 5 questions. A question may be repeated and/or paraphrased only once. If a student still doesn't understand the question, the assessor should move on to the next question.
- Please note that you give **one score** for "Task Achievement (Task One)" based on the students' performance across all 5 questions. You **do not** give a mark per question.

Speaking Assessment Guidelines

ELI
JCEA

Task 2: Short Monologue with Q/A

- The assessors have been provided with 6 pairs of "A" and "B" role cards. The assessors should cycle through these pairs in order. The first two students will be given Prompt# 1 A and Prompt #1 B, the second two students will be given Prompt# 2 A and Prompt # 2B, etc. It is expected that the assessors will cycle through the role 5-6 times depending on the size of their class.
- The interlocutor starts by handing students A and B their respective information cards.
- Students are given **2 minutes** to read their information card, ask questions to the interlocutor, and decide what they want to say. Students are NOT to discuss anything with each other prior to the commencement of the exam.
- After the 2 minutes mentioned above, students begin the task with each student following the directions on his/her prompt in the order mentioned on his/her prompt.
- Students are **individually graded** based on their accomplishment of the bullet points and/or instructions highlighted to them on their card and based on their language use. How well their partner is able to participate should **not** affect the speaker's grade in any way.
- If a student refuses or is unable to participate in the task **or if there is an odd number of students and the last student doesn't have a partner**, the interlocutor must participate as either student A or B as needed. The student not participating should receive a grade of 0 for "Task Achievement (Task Two)". In level 101 and 102 the student would also receive a grade of 0 for "Interaction (Task Two)". For the other rubric categories, the student not participating can receive a **maximum of 3** based solely on his/her responses to Task 1.
  - **Please note:** If the interlocutor is playing the role of Student A or Student B and the prompt tells him/her to talk about something or give his/her opinion about something and then to listen to the other student and ask the other student 2-3 questions, the interlocutor only needs to listen to the other student and ask 2-3 questions. He or she doesn't need to talk about the topic him or herself.

Important points regarding the rubric/assessment rating scale:

- For either task, if the student gives **totally off topic responses, the student should get a zero** for "Task Achievement" for that task. Other categories (like pronunciation, grammar, etc) should be graded **solely** on the basis of the task that wasn't off topic **with a maximum score of 3** for those categories. If **both tasks** have totally off topic responses, then the student should get a zero for the whole exam.
- For either task, if the student gives answers that are **mostly off topic but not completely off** topic, **task achievement should not be rated higher than "2"** and it might be rated "1". The other scores should be rated according to their criteria.
- If a student doesn't meet the requirements for a "1" in any specific rubric category, he/she should be given a "0" for that category.

After all students have been assessed:

- The assessors should total the scores for each student on their "Speaking Exam— PER ASSESSOR Score Recording Sheet".
- Each assessor should then transfer the "total score" column from his "PER ASSESSOR Score Recording Sheet" to the appropriate column in the "Speaking Exam FINAL Score Recording and Attendance Sheet". The classroom teacher is always Assessor 1 on this sheet.
- The classroom teacher should then complete all other sections of the "Speaking Exam FINAL Score Recording and Attendance Sheet" and enter the "Final Score" on the sheet according to the instructions on the sheet. Instructors should **not round** the final mark. They should then enter the final score on the **CA** sheet as well.
- All materials should be returned to the classroom teacher's coordinator.

Appendix B: Speaking Rating Scale in KAU's Preparatory Year English Language Program

102 Speaking Rating Scale

Score	Task Achievement (Task One)	Task Achievement (Task Two)	Interaction (Task Two)	Pronunciation & Fluency
5	Student has sufficient vocabulary to **reasonably accurately** answer **all the** questions using basic sentence patterns with some searching for words.	Student has sufficient vocabulary to **reasonably accurately** tell his/her partner **all the** requested information using basic sentence patterns with some searching for words.	• Student fulfills the requirements of the prompt and can interact **with reasonable ease**, asking clearly understandable questions and effectively answering questions from his/her partner.	• Student's pronunciation is **reasonably accurate** for main vocabulary related to predictable situations. • Can handle short, routine exchanges with **relative ease**. • Hesitations and false starts may still be quite noticeable.
4	Student has sufficient vocabulary to answer **all the** questions using basic sentence patterns with some searching for words. There may be **some** consistent mistakes in basic grammar.	Student has sufficient vocabulary to tell his/her partner **all the** requested information using basic sentence patterns with some searching for words. There may be **some** consistent mistakes in basic grammar.	• Student can answer all of his/her partner's questions **using basic sentence patterns** with **some** consistent mistakes in basic grammar. • Student is able to ask his/her partner comprehensible questions in question formation, with **some** consistent mistakes in basic grammar.	• Student's pronunciation is **reasonably accurate** for main vocabulary related to predictable situations. • Can handle short, routine exchanges with some **difficulty**. • Hesitations and false starts may still be quite noticeable.
3	Student is able to answer **all the** questions using basic sentence patterns, memorized phrases, groups of a few words, and formulae. There may still be **a number of** basic grammar mistakes.	Student is able to tell his/her partner **all the** requested information using basic sentence patterns, memorized phrases, groups of a few words, and formulae. There may still be **a number of** basic grammar mistakes.	• Student can answer all of his/her partner's questions **using basic sentence patterns** with **a number of** consistent mistakes in basic grammar. • Student is able to ask his/her partner comprehensible questions in question formation, with **a number of** consistent mistakes in basic grammar.	• Student can make him/herself understood despite a foreign accent. • **Short utterances**, pauses, false starts, and reformulation **are common.** • Conversational partners may need to ask for repetition from time to time.
2	Student is able to answer **all the** questions **with a very basic repertoire of words and simple phrases.** There may be a number of basic grammar mistakes.	Student is able to tell his/her partner **all the** requested information in the prompt, **using very short, isolated, and pre-packaged utterances with much pausing.**	• Student is able to answer his/her partner's questions **using very short, isolated, and pre-packaged utterances** with much pausing. • Student is able to ask his/her partner questions **but not always in complete question formation, and often with much pausing.** • Student can interact in a simple way but communication is **totally dependent** on repetition, rephrasing, and repair.	• Pronunciation **of a very limited repertoire of learnt words and phrases can be understood with some effort** by native speakers used to dealing with Arabic speakers. • **Short utterances**, pauses, false starts, and reformulation **frequently occur.**
1	Student is **unable to answer 1 or more of the questions with a very basic repertoire of words and simple phrases.**	Student is **unable** to tell his/her partner **1 or more of the things** requested in the prompt, **even in short, isolated, or pre-packaged utterances.**	• Student **may have difficulty** answering one or more of his/her partner's questions. • Student **may have difficulty** asking his/her partner questions that are understandable. • Student **struggles** to interact in a simple way even with the use of repetition, rephrasing, and repair.	• Pronunciation **of a very limited repertoire of learnt words and phrases can be understood with significant effort** by native speakers used to dealing with Arabic speakers. • **Student mostly speaks in short utterances**, with frequent pauses, false starts, and reformulation.
0	Give this score if the student is absent, the language produced is completely irrelevant to the prompts, if the pronunciation is completely unintelligible, **and/or if the criteria for a "1" is not met for the category.**			

Note: Half scores may be awarded (for example, 3.5) for "Pronunciation and Fluency" if the student partially meets the descriptors for 2 bands.

"3" descriptors were based partially on CEFR Table 3 descriptors for A2 Oral Assessment and partially on "St. Giles" A2 performance descriptors (http://www.stgiles-international.com/student-services/level-descriptors); "5" descriptors were based partially on the Jan, 2009 Council of Europe Manual "Relating Language Exams to the CEFR" Table C3 (p.186) for A2+.

Version A

Curriculum and Test Development Unit

Appendix C: Pitch Transcription Conventions (Derived from Pickering, 2009, p. 255)

// //	**tone unit boundaries**
UPPERCASE	prominent syllables
UPPERCASE	tonic syllable carrying the tone choice or tonal pitch movement associated with the tone unit
↑	High pitch
→	Mid pitch

Appendix D: Approval Letter from King Abdulaziz University to Conduct Study

المملكة العربية السعودية
وزارة التعليم
جامعة الملك عبد العزيز
ادارة البعثات

الرقم : ك/40/39515
التاريخ : 1440/09/22هـ

قـــرار إداري
رحلة علمية

الاســـم	: ابراهيم محمد حسين الفيفي	رقـم البطاقة	
الجهة المبتعث اليـها	: جامعة ريدين	رقم السجل المدني	1
الدولـــة	: بريطانيا	رقم ومسمى الوظيفة	
المرحلة الدراسيـة	: الدكتوراه	جهة العمـل	: كلية العلوم والأداب برابغ
التخصـص العام	: اللغه الأنجليزيه وادابها	القسـم	: اللغة الأنجليزية
التخصـص الدقيق	: اللغويات	جهـة الرحلة	: السعـودية
تـاريخ بدايـة الرحلة	: 1440/10/26 هـ 2019/06/29 م	مـدة الرحلة	: ثلاثة أشهر

إن وكيل الجامعة للدراسات العليا والبحث العلمي

* بنـاء على الصلاحيـات الممنوحة له.
* وبناء على توصية اللجنة الدائمة للابتعاث والتدريب رقم (13) المتخذة فى الاجتماع الثاني المنعقد بتاريخ 13/8/1425هـ المتضمن تفويض سعادة وكيل الجامعة للدراسات العليا والبحث العلمي رئيس اللجنة صلاحيات البت فى مواضيع الرحلات العلمية المقدمة من
المبتعثين ليتمكنوا من القيام بها فى الوقت المحدد والمعتمد بخطاب معالي مدير الجامعة رقم 6735/س/25 وتاريخ 27/8/1425هـ
* وبناء على موافقة معالي مدير الجامعة على القيام بالرحلة العلمية نيابة عن مجلس الجامعة بتاريخ 09/09/1440 هـ
* واستنادا إلى المادة (15) من لائحة الابتعاث والتدريب لمنسوبي الجامعات.

تقـــرر ما يلـــي

اولا : يسمح للمبتعث / ابراهيم محمد حسين الفيفي القيام برحلة علمية وفقا للبيـانـات الموضحـة أعلاه .
ثانيا : يكلف أحد أعضاء هيئة التدريس بالقسم التابع له المبتعث بالإشراف الداخلي عليه أثناء قيامه بالرحلة العلمية ومن ثم يرفع تقرير
عما تم إنجازه خلال هذه الرحلة.
ثالثا : يصرف له مخصص إضافي عن كل شهر يقضيه بالمملكة أثناء قيامه بالرحلة العلمية على ألا يزيد ما يصرف له عن مخصص ثلاثة أشهر وذلك بعد تقديم تقريرعن سير الرحلة العلمية معتمد من عميد الكلية.
رابعا : على الجهات المعنية تنفيذ هذا القرار كلا فيمـا يخصـه.

وكيل الجامعة
للدراسات العليا والبحث العلمي

أ.د / يوسف عبد العزيز عبد الله التركي

*صورة لسعادة وكيل وزارة التعليم لشؤون البعثات
*صورة لسعادة الملحق الثقافي السعودي فى بريطانيا (KAU1877)
*صورة لسعادة عميد كلية العلوم والأداب برابغ
*صورة لسعادة مدير عام الشؤون الإدارية والمالية بوزارة التعليم
*صورة لسعادة مدير عام الإدارة العامة لشؤون البعثات بوزارة التعليم
*صورة لإدارةالتخطيط والميزانية والمتابعة/إدارة شئون الموظفين/ المبتعث /ملف المبتعث

Appendix E: Students' Consent Form

School of Literature and Languages
Department of English Language and Applied Linguistics

University of Reading

ETHICS COMMITTEE

Consent Form (Students)

Project title: Nativelikeness and Intelligibility in Pronunciation Assessment from the Perspective of EFL Teachers and Assessment Materials: A Saudi Context

I understand the purpose of this research and understand what is required of me; I have read and understood the Information Sheet relating to this project, which has been explained to me by Ahmad Azzhrani, who is collecting the data on behalf of Ibrahim Alfaifi. I agree to the arrangements described in the Information Sheet in so far as they relate to my participation.

I have received a copy of this Consent Form and of the accompanying Information Sheet.

Name:

Signed:

Date:

3

Appendix F: Students' Information Sheet

University of Reading

Department of English Language and Applied Linguistics

HumSS Building
The University of Reading
Whiteknights, PO Box 218
Reading RG6 6AA

Phone 01183788141
+44 (0)118 378 6472
+44 (0)118 975 6506
Email appling@reading.ac.uk
p.a.thompson@reading.ac.uk

Researcher:
Ibrahim Alfaifi
Email: i.alfaifi@pgr.reading.ac.uk

Supervisor:
Jane Setter
Phone: 0118 3786089
Email: j.e.setter@reading.ac.uk

INFORMATION SHEET (STUDENTS)

Thank you for agreeing to participate in this project, which is undertaken as part of my PhD research at the University of Reading.

You will be asked to read two passages in English, which will be used as audio stimuli in the research project. These recordings will be accessible only to participants in the study, Ahmad Azzhrani (who is collecting the data), myself as PhD researcher and my supervisors, and will be destroyed as soon as the research has been concluded. The recordings will be presented anonymously to the participants. Your privacy and confidentiality will be carefully observed, and all data will be stored on a password-protected computer or in a locked drawer. You can withdraw from participating in the project at any stage.

This project has been subject to ethical review by the School Ethics Committee, and has been allowed to proceed under the exceptions procedure as outlined in paragraph 6 of the University's Notes for Guidance on research ethics.

If you have any queries or wish to clarify anything about the study, please feel free to contact my supervisor at the address above or by email at j.e.setter@reading.ac.uk

Signed

4

Appendix G: Participants' Consent Form

School of Literature and Languages
Department of English Language and Applied Linguistics

University of Reading

ETHICS COMMITTEE

Consent Form (Participants)

Project title: Nativelikeness and Intelligibility in Pronunciation Assessment from the Perspective of EFL Teachers and Assessment Materials: A Saudi Context

I understand the purpose of this research and understand what is required of me; I have read and understood the Information Sheet relating to this project, which has been explained to me by Ibrahim Alfaifi. I agree to the arrangements described in the Information Sheet in so far as they relate to my participation.

I have received a copy of this Consent Form and of the accompanying Information Sheet.

Name:

Signed:

Date:

3

Appendix H: Participants' Information Sheet

University of Reading

Department of English Language and Applied Linguistics

HumSS Building
The University of Reading
Whiteknights, PO Box 218
Reading RG6 6AA

Phone 01183788141
+44 (0)118 378 6472
+44 (0)118 975 6506
Email appling@reading.ac.uk
p.a.thompson@reading.ac.uk

Researcher:
Ibrahim Alfaifi
Email: i.alfaifi@pgr.reading.ac.uk

Supervisor:
Jane Setter
Phone: 0118 3786089
Email: j.e.setter@reading.ac.uk

INFORMATION SHEET (PARTICIPANTS)

Thank you for agreeing to participate in this project, which is undertaken as part of my PhD research at the University of Reading.

You will be asked about your beliefs about what you view as important in pronunciation assessment through three tasks: 1) rating recordings of three EFL Arabic-speaking students each reading two passages; 2) an acceptability task of recorded words/phrases (DMDX task); and 3) a questionnaire. Further, selected participants will then be asked for a follow-up interview to expand their ideas and opinions.

Your responses will be used or accessed only by the PhD researcher and the researcher's PhD supervisor and assessment committee only. Your names will be collected for the research to assist our analysis, but no names or personal details will be revealed in the PhD thesis or subsequent publications. Your privacy and confidentiality will be carefully observed, and all data will be stored on a password-protected computer or in a locked drawer. You can withdraw from participating in the project at any stage.

This project has been subject to ethical review by the School Ethics Committee, and has been allowed to proceed under the exceptions procedure as outlined in paragraph 6 of the University's Notes for Guidance on research ethics.

If you have any queries or wish to clarify anything about the study, please feel free to contact my supervisor at the address above or by email at j.e.setter@reading.ac.uk

Signed

4

Appendix I: The List of the Phonological and Phonetic Features Used in the Assessment of Students' Speech Task

I. Deletion/insertion of sounds in initial consonant clusters (two or more consonants together)
II. Deletion/insertion of sounds in medial consonant clusters
III. Deletion/insertion of sounds in final consonant clusters
IV. The sound /θ/ (e.g. th<u>i</u>ngs)
V. The sound /ð/ (e.g. th<u>e</u>se)
VI. The dark /l/ (e.g., sma<u>l</u>l; in contrast to the light /l/, e.g., <u>l</u>ight)
VII. The sound /r/
VIII. The sound [ɾ] (the American pronunciation of /t/ when it occurs intervocalically (e.g., ma<u>t</u>ter))
IX. The two sounds /nt/ together (e.g., twenty)
X. The aspirated /t/
XI. The aspirated /p/ (e.g., <u>p</u>eas)
XII. The aspirated /k/ (e.g., <u>c</u>all)
XIII. Differences in lengths of vowels preceding final voiced and voiceless consonants (e.g., *b<u>a</u>ck* vs *b<u>a</u>g*)
XIV. Vowels:

- Shortness and length of vowels
- Shortness and length of vowels in diphthongs
- The sound /ɜ:/ (e.g., BrE pronunciation of the underlined sounds in *p<u>ur</u>ple*)

XV. Phrase-level stress (or nuclear stress)
XVI. Intonation

Appendix J: Listening Passages for Assessment of Students' Speech Task

Please call Stella George. Ask her to bring these things with her from the store: six spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a snack for her brother Bob. We also need a small plastic snake and a big toy frog for the twenty kids. She can scoop these things into three yellow or purple bags, and we will go meet her later on Wednesday at the train station.

Before two days, I go to my village because we have there festival for dates, and I have a farm. I take dates from there and they go to festival to pay them because I want money and I want to see how they pay all dates, and this festival the first festival in my village about dates, and all people come from out my village, from Riyadh and Jeddah and Makkah. They come to see how dates about my village, and then we make a party there because people come out my village, and we have dates there. In festival we pay eats there like Kabsa, and all of eat in my village.

My last event I attended was in Jeddah. It is called Jeddah Comic Con. It is event which hold like and target specific audience just like and gamers, and it is multigeneric entertainment, and it is not as international as San Diego Comic Con. It is held in Jeddah, Saudi Arabia, because the place is important in Saudi Arabia, so it must be in Jeddah, and all the people from Dubai and Kuwait come to it because they need a memory of their life because they love the characters and the anime they watch, and they can see also the characters of the anime, and they also can dress like a cosplay which character they love and meet the anime makers stuff like that. Also it is a place where game company show us the game of her like or something. I think that it is all about Comic Con, and it is international so everybody know it.

Before last week, we have a competition in my university. A competition about football game. our university as usual for each years, too much competition, so the first match in our university against department engineering as my departments English language. The first match is very exciting. My team is so-so. Other team is very high, so that match in the beginning in the match we are supporting our department to that game … at the end of the match we win that game and we are going to the last one for the last match for the final match … it is really very exciting because the manager our department it was come and all the supporting and our friends and our professors come that match … it is really exciting that date I can't forget really … and at the begin of that match in the first minutes it is really like rubbish for all my teams really and I can't forget that match and we played extra time for that one … at the end that teams he win for my team.

Appendix K: Assessment of Students' Speech Task

1. **Please listen to the audio recording, and then do the following task: If you find the following items problematic in the speaker's speech, give one example (at least) to them from the passage, otherwise leave them blank.**

I. Deletion/insertion of sounds in initial consonant clusters (two or more consonants together)
Example:

II. Deletion/insertion of sounds in medial consonant clusters
Example:

III. Deletion/insertion of sounds in final consonant clusters
Example:

IV. Mispronunciation of the sound /θ/ (e.g. th<u>th</u>ings)
Example:

V. Mispronunciation of the sound /ð/ (e.g. <u>th</u>ese)
Example:

VI. Mispronunciation of the dark /l/ (e.g., sma<u>l</u>l; in contrast to the light /l/, e.g., <u>l</u>ight)
Example:

VII. Mispronunciation of the retroflex approximant [ɻ] (the American pronunciation of /r/)
Example:

VIII. Mispronunciation of the sound [ɾ] (the American pronunciation of /t/ when it occurs intervocalically (e.g., ma<u>tt</u>er))
Example:

IX. Mispronunciation of the American pronunciation of the intervocalic two sounds /nt/ when they occur before an unstressed syllable (e.g., twenty)
Example:

X. Mispronunciation of the aspirated /t/
Example:

XI. Mispronunciation of the aspirated /p/ (e.g., <u>p</u>eas)
Example:

XII. Mispronunciation of the aspirated /k/ (e.g., <u>c</u>all)
Example:

XIII. Confusion between the lengths of vowels preceding final voiced and voiceless consonants (e.g., *b<u>a</u>ck* vs *b<u>a</u>g*)
Example:

XIV. Confusion between shortness and length of vowels
Example:

XV. Confusion between shortness and length of vowels in diphthongs
Example:

XVI. Mispronunciation of the sound /ɜː/ (e.g., BrE pronunciation of the underlined sounds in *p<u>ur</u>ple*)
Example:

Appendix K (cont.): Assessment of Students' Speech Task

XVII. Misuse of the phrase-level stress (or nuclear stress)
Example:

XVIII. Misuse of intonation
Example:

2. Please provide written responses to the following:

If you find any other phonological or phonetic features problematic in the speaker's speech not mentioned in the list above, provide it below. Also give at least one word example to the features identified from the speaker's speech:

_____
_____
_____
_____
_____
_____.

Appendix L: Demographic Information Questionnaire

Please now fill in the following about yourself.

1. Age: _____
2. First language: _____
3. Other language(s) (Please specify if you are 'native', 'advanced', 'intermediate' or 'beginner' for each language):

   _____

   _____.

4. Qualifications for teaching English:

   _____

   _____.

5. How many years of teaching experience at the university level do you have?

   _____

6. How many times have you been a speaking examiner at the university level?

   _____

7. Do you have any qualifications for assessment of speaking particularly?

   _____

   If yes, please specify: _____

   _____.

8. Have you heard about the Lingua Franca Core? If yes, what do you know about it?

   _____

   _____

   _____

   _____

   _____

Appendix M: The Wordlist Used for the Acceptability Task (It should be noted that an intelligible pronunciation could sometimes be BrE or NAmE pronunciation besides being LFC pronunciation)

No.	Phoneme	Word	IPA transcription	Phoneme position	Type of pronunciation
1.	/θ/	Thick	/t̪ɪk/	Initial	LFC
2.	*	Thumb	/θʌm/	-	BrE/NAmE
3.		Thin	/bɪn/	-	Unacceptable
4.	-	Anything	/enisɪŋ/	Medial	LFC
5.	-	Something	/sʌmθɪŋ/	-	BrE/NAmE
6.	-	Nothing	/nʌmɪŋ/	-	Unacceptable
7.	-	Mouth	/maʊf/	Final	LFC
8.	-	Youth	/juːθ/	-	BrE/NAmE
9.	-	Faith	/feɪd/	-	Unacceptable
10.	/ð/	Than	/d̪æn/	Initial	LFC
11.	-	This	/ðɪs/	-	BrE/NAmE
12.	-	Them	/fem/	-	Unacceptable
13.	-	Within	/wɪvɪn/	Medial	LFC
14.	-	Bathes	/beɪðz/	-	BrE/NAmE
15.	-	Youths	/juːgz/	-	Unacceptable
16.	-	With	/wɪz/	Final	LFC
17.	-	Booth	/buːð/	-	BrE/NAmE
18.	-	Soothe	/suːn/		Unacceptable
19.	/l/	Silk	/sɪ[l]k/	Medial (preceding a consonant)	LFC
20.	-	Milk	/mɪ[ɬ]k/	-	BrE/NAmE
21.	-	Old	/əʊsd/	-	Unacceptable
22.	-	Else	/eʊs/	-	LFC
23.	-	Jelly	/dʒe[ɬ]ɪ/	-	NAmE
24.	-	False	/fɔːns/	-	Unacceptable
25.	-	Channel	/tʃæn[l̩]/	Final (Syllabic)	LFC
26.	-	Middle	/mɪd[ɬ]/	-	BrE/NAmE
27.	-	Stable	/steɪbf/	-	Unacceptable
28.	-	Cycle	/saɪkʊ/	-	LFC
29.	-	Feel	/fiː[ɬ]/	-	BrE/NAmE
30.	-	Apple	/æpæ/	-	Unacceptable
31.	/r/	Run	/[ɻ]ʌn/	Initial	LFC
32.	-	Red	/[ɹ]ed/	-	BrE
33.	-	Rid	/jɪd/	-	Unacceptable
34.	-	Virus	/vaɪ[ɻ]əs/	Medial	LFC
35.	-	Syrup	/sɪ[ɹ]əp/	-	BrE
36.	-	Jerusalem	/dʒəˈvuːsələm/	-	Unacceptable
37.	-	Four eyes	/fɔː[ɻ] aɪz/	Final	LFC

* like the one above.

Appendix M (cont.): The Wordlist Used for the Acceptability Task

No.	Phoneme	Word	IPA transcription	Phoneme position	Type of pronunciation
38.	-	More ice	/mɔː[ɹ] aɪs/	-	BrE
39.	-	Door eye	/dɔːʒ aɪ/	-	Unacceptable
40.	/t/	City	/sɪti/	Medial	LFC
41.	-	Beauty	/ bjuː[ɾ]i/	-	NAmE
42.	-	Lettuce	/ledʒɪs/	-	Unacceptable
43.	/nt/	Ninety	/naɪnti/	Final	LFC
44.	-	Fantasy	/fænəsi/	-	NAmE
45.	-	Bounty	/baʊnzi/	-	Unacceptable
46.	/ɪ/ (epenthesis)	Spaghetti	/sɪpəgeti/	-	LFC
47.	-	Slim	/slɪm/	-	NAmE/BrE
48.	-	Stick	/ˈsɪtɪk/	-	Unacceptable
49.	/ə/ (paragogue)	Pig	/pɪgə/	Final	LFC
50.	-	Dig	/dɪg/	-	NAmE/BrE
51.	-	Big	/bɪgə/	-	Unacceptable
52.	/ɪə/	Mere	/mɪ[˞]/	Final	LFC
53.	-	Here	/hɪə/	-	BrE
54.	-	Beer	/bʊ[˞]/	-	Unacceptable
55.	/ʊə/	Jury	/ dʒʊ[˞]i/	Medial	LFC
56.	-	Lure	/ lʊə/	Final	BrE
57.	-	Ensure	/ɪnʃɪ[˞]/	-	Unacceptable
58.	/eə/	Chair	/tʃe[˞]/	-	LFC
59.	-	Hair	/heə/	-	BrE
60.	-	Wear	/wæ[˞]/	-	Unacceptable
61.	/aʊ/	House	/hɑʊs/	Medial	LFC
62.	-	Mouse	/maʊs/	-	BrE/NAmE
63.	-	Down	/duːn/	-	Unacceptable
64.	/ʊ/	Go	/gʊ/	Final	LFC
65.	-	So	/səʊ/	-	BrE
66.	-	Know	/nʌ/	-	Unacceptable

Appendix N: DMDX Script

<s 1><azk> <ep> <fd 400> <cr> <id keyboard> <nfb><eop>
$0 <@ -5><x 0.2> "Please listen to the following words, and decide", <@ -4><x 0.2> "if their pronunciations are acceptable. "For each item, you will press the RIGHT SHIFT button to answer 'Acceptable'", <@ -1><x 0.2> "pronunciation, or the LEFT SHIFT button to answer 'Unacceptable' pronunciation.", <@ -1><x 0.2> "You have 5 seconds to respond, and then the next word will appear.", <@ -1><x 0.2> "You need to do the task in a single session–you cannot pause. The task is 7 minutes long.", <@ 3> "Press the SPACEBAR to continue.";$
+1 * <wav 2><svp start> "1" /"Thick"/ ;
+2 * <wav 2> <svp start> "2" / "Thumb" / ;
-3 * <wav 2> <svp start> "3" / "Thin" / ;
+4 * <wav 2><svp start> "4" / "Anything"/ ;
-5 * <wav 2><svp start> "5" / "Something" / ;
-6 * <wav 2><svp start> "6" / "Nothing" /;
+7 * <wav 2><svp start> "7" / "Mouth" /;
-8 * <wav 2><svp start> "8" / "Youth" /;
-9 * <wav 2><svp start> "9" / "Faith" /;
+10 * <wav 2><svp start> "10" / "Than"/ ;
-11 * <wav 2><svp start> "11" / "This" /;
-12 * <wav 2><svp start> "12" / "Them" /;
+13 * <wav 2><svp start> "13" / "Within" /;
-14 * <wav 2><svp start> "14" / "Bathes" /;
-15 * <wav 2><svp start> "15" / "Youths" /;
+16 * <wav 2><svp start> "16" / "With" /;
-17 * <wav 2><svp start> "17" / "Booth" /;
-18 * <wav 2><svp start> "18" / "Soothe" /;
+19 * <wav 2><svp start> "19" / "Silk" /;
-20 * <wav 2><svp start> "20" / "Milk" /;
-21 * <wav 2><svp start> "21" / "Old" /;
+22 * <wav 2><svp start> "22" / "Else" /;
-23 * <wav 2><svp start> "23" / "Jelly" /;
-24 * <wav 2><svp start> "24" / "False"/ ;
+25 * <wav 2><svp start> "25" / "Channel" /;
-26 * <wav 2><svp start> "26" / "Middle" /;
-27 * <wav 2><svp start> "27" / "Stable" /;
+28 * <wav 2><svp start> "28" / "Cycle" /;
-29 * <wav 2><svp start> "29" / "Feel" /;
-30 * <wav 2><svp start> "30" / "Apple" /;
+31 * <wav 2><svp start> "31" / "Run" / ;
-32 * <wav 2><svp start> "32" / "Red" / ;
-33 * <wav 2><svp start> "33" / "Rid" / ;
+34 * <wav 2><svp start> "34" / "Virus" / ;
-35 * <wav 2><svp start> "35" / "Syrup" /;
-36 * <wav 2><svp start> "36" / "Jerusalem" /;
+37 * <wav 2><svp start> "37" / "Four eyes" /;
-38 * <wav 2><svp start> "38" / "More ice"/ ;

Appendix N (cont.): DMDX Script

-39 * <wav 2><svp start> “39” / “Door eye” / ;
+40 * <wav 2><svp start> “40” / “City” /;
-41 * <wav 2><svp start> “41” / “Beauty” /;
-42 * <wav 2><svp start> “42” / “Lettuce” /;
+43 * <wav 2><svp start> “43” / “Ninety”/ ;
-44 * <wav 2><svp start> “44” / “Fantasy” / ;
-45 * <wav 2><svp start> “45” / “Bounty” / ;
+46 * <wav 2><svp start> “46” / “Spaghetti”/ ;
-47 * <wav 2><svp start> “47” / “Slim” /;
-48 * <wav 2><svp start> “48” / “Stick” /;
+49 * <wav 2><svp start> “49” / “Pig” / ;
-50 * <wav 2><svp start> “50” / “Dig” / ;
-51 * <wav 2><svp start> “51” / “Big” / ;
+52 * <wav 2><svp start> “52” / “Mere”/ ;
-53 * <wav 2><svp start> “53” / “Here” /;
-54 * <wav 2><svp start> “54” / “Beer” / ;
+55 * <wav 2><svp start> “55” / “Jury” /;
-56 * <wav 2><svp start> “56” / “Lure” /;
-57 * <wav 2><svp start> “57” / “Ensure” /;
+58 * <wav 2><svp start> “58” / “Chair”/ ;
-59 * <wav 2><svp start> “59” / “Hair” /;
-60 * <wav 2><svp start> “60” / “Wear” /;
+61 * <wav 2><svp start> “61” / “House” /;
-62 * <wav 2><svp start> “62” / “Mouse” /;
-63 * <wav 2><svp start> “63” / “Down”/ ;
+64 * <wav 2><svp start> “64” / “Go” / ;
-65 * <wav 2><svp start> “65” / “So” / ;
-66 * <wav 2><svp start> “66” / “Know” /;

$0 <@ 0> “End of task. Thank you so much.”, <@ 1> “Press ESCAPE, and then copy your answers into a Word file and email it to the researcher.”;$

Appendix O: Online Questionnaire

Add item

1 Please type in your first and last names. *

Add item

Add item

2 Please indicate to what extent you agree or disagree with the following statements about English pronunciation teaching and (speaking) assessment. *

	Strong agree	Agree	Slightly agree	Neutral	Slightly disagree	Disagree	Strongly disagree
My students' pronunciations should reflect their cultural identities.	☐	☐	☐	☐	☐	☐	☐
My students' pronunciations should sound nativelike (e.g., like a British/American English speaker).	☐	☐	☐	☐	☐	☐	☐
It is unnecessary for my students to sound nativelike in order to be intelligible (understandable).	☐	☐	☐	☐	☐	☐	☐
Students who sound nativelike are more intelligible speakers.	☐	☐	☐	☐	☐	☐	☐
Students who are intelligible (understandable) will get higher marks.	☐	☐	☐	☐	☐	☐	☐
The more students sound nativelike the higher mark they get.	☐	☐	☐	☐	☐	☐	☐
It is good for my students to have an Arabic accent when speaking because they are Arabic speakers.	☐	☐	☐	☐	☐	☐	☐
Non-native speakers are valid pronunciation models.	☐	☐	☐	☐	☐	☐	☐
Sounding nativelike is more important than being intelligible.	☐	☐	☐	☐	☐	☐	☐
Classroom materials for pronunciation should include nativelike models of English rather than other non-native ones.	☐	☐	☐	☐	☐	☐	☐
It is important that students be exposed to English used by intelligible (understandable) non-native speakers.	☐	☐	☐	☐	☐	☐	☐
I want to teach my students to pronounce English like native speakers.	☐	☐	☐	☐	☐	☐	☐
I want to teach my students to be intelligible (understandable) in their spoken English.	☐	☐	☐	☐	☐	☐	☐
(For non-native speakers only): I would feel proud if my students perceive my accent as a native accent of English.	☐	☐	☐	☐	☐	☐	☐
(Native speaker teachers only): I am proud when students comment favourably on my native speaker accent.	☐	☐	☐	☐	☐	☐	☐

Add item

Appendix P: Intelligibility and Nativeness Statements in the Online Questionnaire

Intelligibility statements:

- My students' pronunciations should reflect their cultural identities.
- It is unnecessary for my students to sound nativelike in order to be intelligible (understandable)
- Students who are intelligible (understandable) will get higher marks.
- It is good for my students to have an Arabic accent when speaking because they are Arabic speakers.
- Non-native speakers are valid pronunciation models.
- It is important that students be exposed to English used by intelligible (understandable) non-native speakers.
- I want to teach my students to be intelligible (understandable) in their spoken English.

Nativeness statements:

- My students' pronunciations should sound nativelike (e.g., like a British/American English speaker).
- Students who sound nativelike are more intelligible speakers.
- The more students sound nativelike the higher mark they get.
- Sounding nativelike is more important than being intelligible.
- Classroom materials for pronunciation should include nativelike models of English rather than other non-native ones.
- I want to teach my students to pronounce English like native speakers.
- (For non-native speakers only): I would feel proud if my students perceive my accent as a native accent of English/ (For native speaker teachers only): I am proud when students comment favourably on my native speaker accent.

Appendix Q: The List of the Follow-up Questions in the Follow-up Interview

- Do think they were appropriate for measuring teachers' attitudes to intelligibility and nativeness?
- I mean any background factors you think have affected your responses in the tasks such as particular experience, previous speaking assessments?
- Anything you want to add to your responses to the previous questions?
- Do you think they were good for measuring attitudes to intelligibility and nativeness in pronunciation assessment?
- I see. But what is the reason for this conflict I found? Do you think of any specific reason?
- Can you think of any specific ones?
- Do you think speaking assessments where you work or where you've worked had any effect?
- All the tasks were used to find out if you were oriented to intelligibility or nativeness. There is a framework called "the LFC" that was used to measure if you were oriented to either constructs. Based on your responses to the tasks, it seems you were more oriented to intelligibility in the online questionnaire but oriented to nativeness in the two practical tasks. So .. can you tell of any reasons behind such conflict?
- Like speaking assessments where you work .. do you think they had any effect on your attitudes in the tasks?
- Yes. What do you think of the listening tasks and the questionnaire?
- What about speaking assessments where you work? Do you think they had any effect on your attitudes in the tasks?
- I sent you a while ago two assessment tasks. One of them was rating speakers on a list of phonetic and phonological features, and the other one using a programme you install and then run the task on it. You also did for me a questionnaire online. Do you think these tasks were appropriate for eliciting teachers' attitudes regarding intelligibility and nativeness.
- They were looking at the same thing. But what do you think why the tasks showed you were oriented to nativeness but the online questionnaire showed you were oriented to intelligibility?
- No. What do you think of them as tools for finding attitudes regarding intelligibility and nativeness? Do you think they were successful in eliciting these attitudes about EFL teachers?
- I mean any background factors that affected your attitudes and behaviour while doing the tasks and the online questionnaire. Do any factors such as the speaking assessment polices at King Abdulaziz University had any effect?

Appendix R: Transcription Conventions Used for the Interviews of this Study (Adopted from Schiffrin, 1994, pp. 422-433)

Speakers	Symbol
Speaker identity/turn start	:
Speech overlap	[

Transitional continuity	Symbol
Final	.
Continuing	,
Appeal	?

Speech factors	Symbol
Continuing/continued speech	➜
Pause (long)	…
Pause (short)	..

Appendix S: Transcription of the Follow-up Interviews

Dialog with Interviewee 1:

Total Time: 00:15:35

Interviewer: You have done three tasks for my study: an assessment of students' speech task, an acceptability task, and an online questionnaire. What do you think of these tasks?

Interviewee: I think they are good .. Actually, I don't remember a lot, but from which aspect you mean?

Interviewer: Do think they were appropriate for measuring teachers' attitudes to intelligibility and nativeness?

Interviewee: The online questionnaire I remember it was asking about this, but the other tasks I was sure what their purpose.

Interviewer: Okay, never mind. Your behaviour in the three tasks you have done show that you were oriented to intelligibility in the online questionnaire but oriented to nativeness in the two practical tasks: the assessment of students' speech and acceptability tasks. Why?

Interviewee: I don't know .. but I think intelligibility is the benchmark for speech performance. So .. I think I answered the questionnaire based on this .. in the other tasks .. I think I was looking for intelligible speech.

Interviewer: What were the factors that affected your responses to the tasks?

Interviewee: Like what?

Interviewer: I mean any background factors you think have affected your responses in the tasks such as particular experience, previous speaking assessments?

Interviewee: Sorry but I really can't think of any factors such as these.

Interviewer: Do you have anything else you want to add or say?

Interviewee: I'm not sure .. like what?

Interviewer: Anything you want to add to your responses to the previous questions?

Interviewee: Actually no.

Interviewer: Okay, no problem. Thanks so much.

Appendix S (cont.): Transcription of the Follow-up Interviews

Dialog with Interviewee 2:

Total Time: 00:15:58

Interviewer: You have done three tasks for my study: an assessment of students' speech task, an acceptability task, and an online questionnaire. What do you think of these tasks?

Interviewee: In which sense?

Interviewer: Do you think they were good for measuring attitudes to intelligibility and nativeness in pronunciation assessment?

Interviewee: I am not sure, but in general yeah they were fine.

Interviewer: Okay. Your behaviour in the three tasks you have done show that you were oriented to intelligibility in the online questionnaire but oriented to nativeness in the two practical tasks: the assessment of students' speech and acceptability tasks. Why?

Interviewee: My attitudes should be the same everywhere. I am surprised you found something different.

Interviewer: I see. But what is the reason for this conflict I found? Do you think of any specific reason?

Interviewee: Sorry, I really don't have an answer for your question. I told you I was thinking my attitudes should be the same. It is just you found something different. But I think intelligibility was the thing I was looking for in the tasks. I think I prefer intelligibility to nativeness and my answers were based on this.

Interviewer: Never mind. Now what were the factors that affected your responses to the tasks?

Interviewee: There could be many.

Interviewer: Can you think of any specific ones?

Interviewee: I really don't have an answer for this question. It may, however, be due to my lack of awareness of the terminology. I had to look up the terms "nativeness" and "intelligibility" principles back then.

Interviewer: Do you think speaking assessments where you work or where you've worked had any effect?

Interviewee: No, I don't think so.

Interviewer: Do you have anything else you want to add or say?

Interviewee: No. Thank you for having me.

Interviewer: Thank you for this interview and for all your help.

Appendix S (cont.): Transcription of the Follow-up Interviews

Dialog with Interviewee 3:

Total Time: 00:16:07

Interviewer: You have done three tasks for my study: an assessment of students' speech task, an acceptability task, and an online questionnaire. What do you think of these tasks?

Interviewee: They took me a long time, but I think they were clear and easy to do. It was just the other task that needs a programme to run .. it was a bit tricky to install the programme and run it. But overall all is fine and well-structured.

Interviewer: Nice to hear that. Your behaviour in the three tasks you have done show that you were oriented to intelligibility in the online questionnaire but oriented to nativeness in the two practical tasks: the assessment of students' speech and acceptability tasks. Why?

Interviewee: Sorry, I just don't get it .. how my attitudes were different? I mean how did you find out?

Interviewer: All the tasks were used to find out if you were oriented to intelligibility or nativeness. There is a framework called "the LFC" that was used to measure if you were oriented to either constructs. Based on your responses to the tasks, it seems you were more oriented to intelligibility in the online questionnaire but oriented to nativeness in the two practical tasks. So .. can you tell of any reasons behind such conflict?

Interviewee: I believe that .. nativeness and intelligibility are somehow interrelated. Being able to speak nativelike mandates being intelligible. So .. focusing on nativelike pronunciation when teaching students, for instance, .. will necessarily lead to intelligibility.

Interviewer: I see. Now what were the factors that affected your responses to the tasks?

Interviewee: Can you give any examples of such factors?

Interviewer: Like speaking assessments where you work .. do you think they had any effect on your attitudes in the tasks?

Interviewee: I think yeah, following the policies in my work must have formed some of my opinions.

Interviewer: Do you have anything else you want to add or say?

Interviewee: I think I've said what I meant to say regarding your questions. So .. no thanks I have nothing more to say than thank you.

Interviewer: Thanks to you, too.

Appendix S (cont.): Transcription of the Follow-up Interviews

Dialog with Interviewee 4:

Total Time: 00:16:16

Interviewer: You have done three tasks for my study: an assessment of students' speech task, an acceptability task, and an online questionnaire. What do you think of these tasks?

Interviewee: As far as I remember they were .. like two listening tasks and a questionnaire I did online.

Interviewer: Yes. What do you think of the listening tasks and the questionnaire?

Interviewee: Yeah, they were like two listening tasks .. one of them was asking to rate speakers, and there was like a list of things to rate them on. Yeah, I think the list was like a kind of difficult to use to rate the speakers. I think because I'm not familiar with such kinds of tests. I'm not sure.

Interviewer: OK, one more question. Your behaviour in the three tasks you have done show that you were oriented to intelligibility in the online questionnaire but oriented to nativeness in the two practical tasks: the assessment of students' speech and acceptability tasks. Why?

Interviewee: I'm not sure how to spot on that, but my understanding my attitudes must be the same. So, I'm not sure why and how my attitudes were different to each other.

Interviewer: I understand. Okay, what were the factors that affected your responses to the tasks?

Interviewee: Maybe the education I had in school affected my opinions. When I was a student, especially at university, the teachers, the textbooks, the audios for listening .. all of these taught the British accent.

Interviewer: What about speaking assessments where you work? Do you think they had any effect on your attitudes in the tasks?

Interviewee: I don't think so. I think the learning experience is the cornerstone or .. the basis for our attitudes as teachers. There is a saying in Arabic .. learning during childhood is like engraving in a stone. So, it's something removable. So, what we learn during childhood, I think it's like engraves attitudes in ourselves.

Interviewer: Do you have anything else you want to add or say?

Interviewee: It's just about the listening tasks. I still don't believe that they can tell our attitudes. Sorry, but this is my understanding, but I think you know what you're doing.

Interviewer: I see. Thanks so much for participating.

Appendix S (cont.): Transcription of the Follow-up Interviews

Dialog with Interviewee 5:

Total Time: 00:15:54

Interviewer: You have done three tasks for my study: an assessment of students' speech task, an acceptability task, and an online questionnaire. What do you think of these tasks?

Interviewee: Sorry, what do you mean?

Interviewer: I sent you a while ago two assessment tasks. One of them was rating speakers on a list of phonetic and phonological features, and the other one using a programme you install and then run the task on it. You also did for me a questionnaire online. Do you think these tasks were appropriate for eliciting teachers' attitudes regarding intelligibility and nativeness.

Interview: Yeah, the online questionnaire was asking about that, I remember. But the other tests I didn't realise they were for the same purpose.

Interviewer: OK. Next. Your behaviour in the three tasks you have done show that you were oriented to intelligibility in the online questionnaire but oriented to nativeness in the two practical tasks: the assessment of students' speech and acceptability tasks. Why?

Interviewee: I believe intelligibility should be more important than sounding like a native speaker. My answers in the questionnaire, I think, showing my opinion. I mean showing that intelligibility is more important. The other tasks, I didn't know they were checking the same thing.

Interviewer: They were looking at the same thing. But what do think why the tasks showed you were oriented to nativeness but the online questionnaire showed you were oriented to intelligibility?

Interviewee: I don't know why. But my belief is that intelligibility is more important.

Interviewer: I understand. Now, what were the factors that affected your responses to the tasks?

Interviewee: Maybe because my learning experience .. all the things I learned was in American English

Interviewer: Okay, but what about the speaking assessments where you teach now? Do they affect your responses in the tasks?

Interviewee: No, they don't.

Interviewer: Do you have anything else you want to add or say?

Interviewee: No, Thanks.

Interview: Thank you for your time.

Appendix S (cont.): Transcription of the Follow-up Interviews

Dialog with Interviewee 6:

Total Time: 00:15:46

Interviewer: You have done three tasks for my study: an assessment of students' speech task, an acceptability task, and an online questionnaire. What do you think of these tasks?

Interviewee: Yeah, they were good. What do you mean? I mean I should tell you how good they were?

Interviewer: No. What do you think of them as tools for finding attitudes regarding intelligibility and nativeness? Do you think they were successful in eliciting these attitudes about EFL teachers?

Interviewee: They could be successful. Maybe I need to go back to the tasks to check that clearly.

Interviewer: Your behaviour in the three tasks you have done show that you were oriented to intelligibility in the online questionnaire but oriented to nativeness in the two practical tasks: the assessment of students' speech and acceptability tasks. Why?

Interviewee: Yeah, it's clearly I'm oriented to intelligibility. In the questionnaire I was asked to state my opinion about statements on intelligibility, and I think I clearly stated my position. In the tasks, it was not clear to me they were asking me to state my position regarding intelligibility. But I'm sure I rated the listenings recorded according to their intelligibility. I really don't know why you found my attitudes were in conflict in this regard.

Interviewer: Okay, I understand. What were the factors that affected your responses to the tasks?

Interviewee: Factors like what?

Interviewer: I mean any background factors that affected your attitudes and behaviour while doing the tasks and the online questionnaire. Do any factors such as the speaking assessment polices at King Abdulaziz University had any effect?

Interviewee: I think I was affected by these assessments .. , and also the assessments when I was teaching in my country.

Interviewer: Okay. Now, do you have anything else you want to add or say?

Interviewee: I can't think of anything more to say. Maybe you can ask me about specific things to help me talk more.

Interviewer: That's it. Thanks so much

Appendix T: A Summary of the Output of the Correlation Analyses for the Dependent Variables of the Practical Tasks

		Intelligibility responses of assessment of students' speech task	Nativeness responses of assessment of students' speech task	Intelligibility responses of acceptability task	Nativeness responses of acceptability task
Intelligibility responses of assessment of students' speech task	Correlation Coefficient	---	.105	.220	.191
	Sig. (2-tailed)	---	.437	.099	.156
Nativeness responses of assessment of students' speech task	Correlation Coefficient	.105	---	-.135	-.143
	Sig. (2-tailed)	.437	---	.316	.289
Intelligibility responses of acceptability task	Correlation Coefficient	.220	-.135	---	.134
	Sig. (2-tailed)	.099	.316	---	.320
Nativeness responses of acceptability task	Correlation Coefficient	.191	-.143	.134	---
	Sig. (2-tailed)	.156	.289	.320	---

*p<0.05, **p<0.01, ***p<0.001