



Universiteit Leiden

Second language users' attitudes towards the 'correctness' of global English pronunciations

Amy Catling

S2002388

MA Linguistics, Language and Communication

June 2018, Universiteit Leiden

Academic Supervisor: d.smakman@hum.leidenuniv.nl

Second Reader: j.grijzenhout@hum.leidenuniv.nl

Abstract:

In a pronunciation evaluation task, 30 Dutch students of English Language & Culture responded to questions about the 'correctness' in pronunciation of 4 speakers of global Englishes. This included two L1 speakers (from London and New Delhi), and two L2 speakers (from Groningen and Zhengzhou). Follow-up interviews were conducted with 10 participants for further insight. Results showed that the pronunciations were ranked as most to least correct in regard to their similarity to RP. However, three factors were highly influential in participants' evaluations. Firstly, prosodic features were found to be the defining feature of 'good' pronunciation. Secondly, identification of a speaker's accent affected participants' overall evaluations. Finally, a speaker was thought to have a lower standard of pronunciation if they did not have an RP accent, regardless of how easy they were to understand. Overall, English was found to be indexical of the UK or USA due to RP and GA being the standard pronunciation models, and if these teaching models are to continue to be the norm, there is an overwhelming need to increase the use of other pronunciation models as stimuli in the classroom to address the practical realities of using English today.

Table of Contents

1. Introduction	3
1.1 Background	3
1.2 Research questions	4
2. Literature Review	6
2.1 The future of global English pronunciation	6
2.2 The preferred English pronunciation of second language learners	7
2.3 Attitudes towards ‘foreign’ accented English.....	8
2.3.1 <i>Native speaker attitudes towards non-native speech</i>	9
2.3.2 <i>Non-native speaker attitudes towards non-native speech</i>	10
3. Pronunciation Evaluation Task	12
3.1 Method	12
3.1.1 <i>Participants</i>	12
3.1.2 <i>Stimuli</i>	12
<i>The Four Speakers</i>	13
<i>The Recording Material</i>	14
3.1.3 <i>Procedure</i>	16
<i>Background Questions</i>	17
<i>Evaluation Tasks</i>	17
3.2 Results	18
3.2.1 <i>The overall classification of speakers</i>	18
3.2.2 <i>Ranking features of pronunciation</i>	19
<i>General trends for the ranking of pronunciation features on Likert scales</i>	19
<i>General trends for the individual scores of speakers</i>	20
3.2.3 <i>Determining which features impact pronunciation the most</i>	21
3.2.4 <i>Identification of the pronunciation</i>	24
3.2.5 <i>Association of RP as being English</i>	26
3.2.6 <i>Comprehensibility and native speaker scores</i>	27
4. Follow-up Interviews	30
4.1 Method	30
4.1.1 <i>Participants</i>	30
4.1.2 <i>Procedure</i>	30
4.2 Results.....	31
4.2.1 <i>Being understandable vs. sounding native-like</i>	31
4.2.2 <i>Attitudes on the native speaker model</i>	32
4.2.3 <i>Experience with Dutch second language speakers</i>	34
5. Discussion	36
5.1 Overview	36
5.2 The defining feature of good pronunciation	36
5.3 The familiarity of the pronunciations	38
5.3.1 <i>Familiar pronunciations</i>	38
5.3.2 <i>Unfamiliar pronunciations</i>	39
5.4 Understanding pronunciation and the ideology behind the native speaker model	40

5.4.1 Implications on English teaching.....	41
6. Conclusion.....	43
6.1 Overall findings.....	43
6.2 Implications for future research.....	44
References.....	46
Appendices.....	50

1 Introduction

1.1 Background

English language use around the world involves a highly multicultural and diverse range of people. There are now more non-native speakers of English in the world than native (Galloway & Rose 2015). Although Trudgill (2005:78) argues that the majority of English usage is still by native speakers, it is clear that English continues to expand as the current international lingua franca, with a growing base of second language users (Rogerson-Revell 2014; Jenkins 2007; Canagarajah 2007; Sung 2014).

The variety of different first language backgrounds among current English speakers has influenced and caused there to be a large number of spoken accents, or global Englishes. Despite this known diversity, other speakers often judge a person's linguistic ability, identity and status through pronunciation. In these cases, a defined standard set of pronunciation ideals is used to define some English as 'good,' meaning that others must therefore be 'bad' (Lippi-Green 1997).

The language ideology that there is a standard spoken form of English means that few variations are considered indexical of 'good' pronunciation. The two widely accepted varieties of English for second language teaching are Received Pronunciation (RP), often associated with southern UK speakers, and General American (GA), often associated with Midwestern USA speakers (Galloway & Rose. 2015). Other forms of speech are therefore seen as non-standard, or even sub-standard. Pronunciation similar to RP or GA is often defined as a 'neutral' or 'correct,' whereas systematic non-standard pronunciation is defined as accented speech - despite the fact RP and GA are accents themselves (Lippi-Green 1997:72).

Pronunciation cannot be isolated from other aspects of language; comprehension of semantic and pragmatic cues is still necessary to gauge meaning in a person's second language. Despite this, research has shown that second language accented speech is often regarded negatively by both first language and second language users alike (Dragojevic et al 2017; Hu & Lindemann 2009; McKenzie 2008). Other studies have shown that transfer errors from a speaker's first language are almost impossible to completely eradicate (Szpyra-Kozłowska 2015:7). It is interesting to note that within the same study, Derwing & Munro (2009) state that most students will be unlikely to ever fully lose the influence from their first language on their second, yet maintain that accent reduction is important for communicative effectiveness, highlighting a continued emphasis on the native speaker model. But with native speakers ranging from Scottish to Singaporean, and with equally proficient non-native English speakers originating from places as diverse as Sweden and Brazil, why are some spoken global varieties considered more synonymous with speaking 'correct English' than others?

Hendriks et al (2017:58) have suggested that strongly accented English is much more likely to lead to negative evaluations than mildly accented English. However, in order to determine the specific features that cause a specific positive or negative evaluation, it is important that evaluation tasks

maintain a consistently high level of intelligibility for different international accents. This will help avoid interference in the results due to potential incomprehensibility of strong, unfamiliar accents.

This study intends to investigate the specific phonetic or sociolinguistic features which prompt a positive or negative evaluation on the 'correctness' of 'mild' international Englishes. The international Englishes used represent speakers from each of Kachru's (2009) inner, outer (both first language speakers) and expanding circles (second language speakers) of English. The 'mild' category of accentedness is defined in this study by a speaker reaching Colin et al's (2011) third level in the hierarchy of error, meaning that all speakers must have high intelligibility, low irritation or distracting transfer errors, yet have a detectable accent. Using this criteria for all accents evaluated means that the study can attempt to explain the reasons why certain positive and negative evaluations occur about different international English accents, in a context in which comprehensibility of the speaker should not be called into question.

In particular, this study will investigate the views on the 'correctness' of international Englishes by Dutch students majoring in English Language & Culture at Universiteit Leiden; all of whom have a background in English pronunciation training. These results will show whether the current second language education system of English emphasises the acquisition of native-like pronunciation in order to be regarded as 'good' at speaking English. Furthermore, the people in this group hope to become English teachers, and may be the people who classify the future English-speaking norms of the Netherlands. The English Proficiency Index (EFI 2018), reported in 2017 that the Netherlands had the highest second language English proficiency in the world, with average national proficiency test scores of 71.45%. With such high standards of English education, the Netherlands may set the future standard for second language education worldwide. Therefore, these students' thoughts on 'correct,' or in their opinion, 'good,' pronunciation may also bring insight into the future of English language teaching, and whether exposure and teaching of global English pronunciations are likely to be of importance in the second language classroom.

1.2 Research Questions

My research questions for this investigation can be summarised as follows:

R1: *Do Dutch students majoring in an English Language & Culture BA evaluate first language speakers' pronunciation of English more positively than second language speakers' pronunciation of English?*

R2: *What are the most important factors in determining either a perceived 'correctness' or 'incorrectness' of English pronunciation?*

To address these questions, this study involves two experiments. Firstly, a pronunciation evaluation task, in which Dutch students of English will be asked to evaluate and comment on pronunciation features of 4 highly intelligible yet accented speakers; including 2 first language speakers from both the inner and outer circles of English, and 2 second language speakers from the expanding circle of

English. Secondly, a series of short follow-up interviews in which other students majoring in the same course will be asked about their definition of 'correct,' or 'good,' English pronunciation, without any stimuli to prompt their views. Using the answers, I will also consider how the participants' evaluations could impact future second language English pronunciation teaching.

Based on the consistencies in previous literature that state second language pronunciation with interference from a first language is most likely to be deemed negatively (Dragojevic et al 2017; Lippi-Green 1997), and English speakers are most often associated with an Anglo-American background (Galloway & Rose. 2015), my hypotheses are the following:

H1: Dutch students majoring in English Language & Culture will regard the speakers with the least amount of deviation from the standard taught norms of RP or GA as being the most correct. As the second language users are more likely to have interference, these will be evaluated more negatively than the first language speakers.

H2: The association of RP and GA being indexical of English will strongly impact the students' evaluations of the pronunciations, resulting in the most important factors in identifying 'correct' and positively viewed English pronunciation as being determined by how similar the speaker's pronunciation is to RP or GA.

2 Literature Review

2.1 The future of global English pronunciation

The starting point of this section is to consider the possible futures of English pronunciation teaching norms to address its international population of users, and see how realistic this is from its current standing point.

Recent studies have suggested that because most English speakers are non-native, the majority of English interactions will not include a native speaker (Canagarajah 2007:925). Subsequent research has examined the role of English as the current international lingua franca (ELF) - and whether ELF should be taught with the primary objective of being understood internationally, differing from traditional teaching of English as a second language (Seidlhofer 2011; Cogo 2012). A major principle of ELF theory is that aiming to replicate a restricted view of native-speaker talk is unrealistic and does not represent the modern status of English as a global language. Among other linguistic differences, this worldwide variation of Englishes leads to differences in pronunciation, and ELF research aims to bridge the gap such that if a linguistic item is produced in a different way than the standard RP or GA, 'it is not automatically an error.' (Jenkins 2009:141)

This idea has caused much controversy - receiving both support (Jenkins 2009; Seidlhofer 2011), and questions about its practical use (Sowden 2012; Sewell 2013). Despite the initial view of Jenkins (2009) that all users of English should have a say in how the language is used, there appear to be some limitations within her argument. She proposes a list of ELF pronunciation features that could set the norm for future ELF contexts, including eradicating the need for the phonemes such as /θ/ and /ð/ and the teaching of the fortis-lenis distinction, which are prone to common transfer errors. However, these suggestions appear to be based on little empirical research (Haslam & Zetterholm 2016). Furthermore, English is presented as the chosen lingua franca language when speakers have no other mutual language, yet code-switching is mentioned as a possibility to help ELF exchange. Code-switching cannot have a place unless both parties have some command of both languages, and if this is the case, it is not clear why English would be chosen as the lingua franca when the purpose of ELF is to guide communication in the simplest way possible.

Seidlhofer (2011:24) states that ELF theory needs to be empirically tested and argues that ELF descriptions of forms should be recorded to prove that the concept is not just a simplified version of English as a second language. However, the danger with using ELF research to create an international set of English pronunciation norms for the future is that one set of pronunciation norms may be replaced by another set of norms, and that it is impossible to accommodate all first language backgrounds to the advantage of all speakers.

The true goal of ELF research should be to interpret how multiple varieties of English communicate effectively despite differences in pronunciation. By identifying the pronunciation features that help or impede communicative effectiveness for speakers from different linguistic backgrounds, these findings would help new learners to develop techniques to adapt to speakers with a variety of

pronunciations. I agree with Seidlhofer (2011:116) that within language change and ‘the ideologies and beliefs associated with the language,’ the fact that there are more second language users of English than first indicates that both native and non-native groups should have a say in the future of the language. Language change is inevitable, and the future pronunciation norms of the English language are most likely to be affected and brought on by the diversity of its speakers. However, I disagree that ELF norms should be set by outside research, as this attempted control of language change does not reflect the everyday uses of English, just as the current pronunciation norms in teaching do not reflect the majority of English speakers.

Looking at ELF in practice and the opinions of its speakers, Sung (2014:53) concluded that identity had a large role to play in whether someone would want to adhere to certain pronunciation preferences or not. This is consistent with other studies that claim identity can strongly impact a person’s choice of L2 accent (Cutler 2014; Kirkpatrick 2007). Within Sung (2014), 9 Hong Kong students had contrasting opinions about using an Anglo-American pronunciation when speaking English that was highly dependent on whether they felt it made them a part of the ‘global community’ or whether they preferred to retain their local identity in speech.

The value of identity and people’s wish to have a choice in how they represent themselves is important to consider - even within a lingua franca context where communication is the primary objective. The role of English as a lingua franca is undeniable. But its future pronunciation norms are difficult to predict because there are personal, social, and linguistic factors that influence each speaker’s choice of norms. To investigate language, we also need to investigate the users of that language.

With mixed feelings surrounding the international nature of English and whether certain pronunciation norms should be used in teaching the language, it is important to consider the needs and attitudes of the students who are in the process of learning it (Friedrich 2000; Starks & Paltridge 1996). The next section will examine literature about the preferred pronunciation norms students reportedly want to learn.

2.2 The preferred English pronunciation of second language learners

Pronunciation training may not always be an explicit part of the second language curriculum. However, pronunciation norms are often implied in the classroom. Even with the knowledge that there are many varieties of systematic English pronunciation norms, when learning it as a second language, it is often the case that language learning materials become synonymous with one variety, and one type of people (Galloway & Rose 2015:197).

When asking students about their preferred choice of English pronunciation to be taught, Abeywickrama (2013) and Szpyra-Kozłowska (2015) found that the majority of participants (65% and 73%, respectively) chose either GA or RP. In both cases, this choice was thought to be connected to a belief that ‘American and British English’ are more easily understood internationally, and learning a non-native variety would hamper students’ language use. Furthermore, in Szpyra-Kozłowska (2015),

some participants gave reasons for choosing these varieties such as cultural and historical interest in either the UK or USA, deciding against an 'international' pronunciation due to its lack of cultural association. Students often saw learning English as an opportunity to connect with one of these two personas, and the intention of using English as a lingua franca was often secondary.

Despite this finding, when Abeywickrama (2013) examined how the same students responded to a comprehension test that included both native and non-native speakers of English, results showed there was no significant difference in the intelligibility rating of any speaker, and 2 of the 4 US teaching assistants who provided recordings were incorrectly identified as being non-native. Even though these students felt sounding like a native speaker of English is the preferred option, in practice, they were not always able to recognise who these speakers are. With this in mind, the main conclusion we can take from the above studies is that these students associate 'native' English (associated as being RP or GA) with 'good' pronunciation. To take this a step further, these students may choose these varieties as the best option to be taught, to prove their own English proficiency and sound 'good' themselves.

Do these preferences come from the established nature of RP and GA in the classroom, or the external representation of the English language? Empirical research on US media has shown that GA dominates television and cinema, meaning the diversity of English varieties in the US is not accurately represented (Lippi-Green 1997; Dragojevic et al 2016). Within a classroom setting, van den Doel (2006:3) stated that there is 'overemphasis in second language acquisition on the standard language,' as the native English population is far from homogeneous. In an analysis of the Common European Framework of Reference (CEFR) for English language teaching, Pitzl (2015:99) found that although the native speaker model was explicitly rejected, within the proposed goals for learners of English, the 'benchmarks' of language learning correlate to replicating particular native speaker standards. In other words, the English language is often externally synonymous to second language learners with these two particular varieties through both the curriculum and the media. With this in mind, any native deviation from this established norm is classed as 'non-standard' - and therefore may be classed as being 'wrong.' Thus, students' pronunciation preferences may be influenced by a pre-established 'correct' way of reflecting the English 'cultural identity.' Asking for a change in what they are taught and shown may be seen as asking to be taught variants that are not 'correct' forms - and could internationally be evaluated as not being 'good' English (Geeslin & Long. 2014:258).

However, the above studies discuss pronunciation preference at an ideological level, considering what students think they should be taught, rather than reflecting how they feel about different English pronunciations used in practice. The next section will examine how second language pronunciation of English is viewed in practice, outside of a classroom setting.

2.3 Attitudes towards 'foreign' accented English

Some researchers have opted to describe second language pronunciation of English as 'foreign' accented English, with foreign accents defined as differing in either (or both) its segmental and prosodic features from a standard variety (Behram 2014:547). I disagree with this terminology as

'foreign' suggests a type of pronunciation that is unusual or distinct from what it 'should be,' with no regards of whether this term is exclusive to second language users only. However, in discussing the below literature I will use the term 'foreign accent' to be consistent with the terminology used in the studies themselves.

Many studies have shown that there are, generally, negative attitudes towards 'foreign' accented English speech (Dragojevic et al 2017; Hu & Lindemann 2009; McKenzie 2008). Within this section, the trends identified in these studies will be highlighted, and will lead to a discussion on how this study will be addressing a gap in the literature.

2.3.1 Native speaker attitudes towards non-native speech

In matched-guise tests that examine native speakers of English's views on non-native Englishes, results have shown that the higher the perceived 'foreignness' of a speaker, the lower the evaluation of the accent, and leading on from that, the lower the reported intelligibility of it. This has been the case for participants of South African English (Coetzee-Van Rooy 2009), American English (Dragojevic et al 2017; Lindemann 2003), and Australian English participants (Fraser & Kelly 2012), to name a few. These results reinforce Kachru's (2009:284) 'interlanguage myth' that describes non-native language varieties as being forms of incomplete communication not yet achieving their goal of sounding native-like, as the intelligibility of a heavily accented non-native speaker is questioned by participants. This has led to suggestions that a listener will evaluate a strong non-native accent as being harder to process, and this will affect their end perceptions of their pronunciation, regardless of how intelligible or not it is in reality (Dragojevic et al. 2017:391).

Dragojevic & Giles (2016) tested the intelligibility rating by US participants of two native speakers of English with different perceptions of 'foreignness': one GA and one Punjabi English speaker, in 4 recordings. Each speaker made one recording which was used twice: once with added white noise, and a second left without. This methodology intended to observe whether the higher difficulty in processing the recording with white noise disturbance would affect the reported intelligibility rating for both speakers, regardless of their English pronunciation. The results showed no significant difference in intelligibility rating in either condition for the GA speaker, whereas the Punjabi English speaker was found to be much less intelligible when white noise was added. In this case, has the difficulty of processing the more 'foreign' speaker with white noise been interpreted as fault of the speaker because their pronunciation is not as familiar to the participants?

The familiarity of the accent to a listener is a consistent secondary theme within the above studies. Lindemann's (2003) study on the perceptions of Korean speakers found that native-US participants who regularly communicated with Korean Americans in their community of practice responded to a Korean English accent more positively. Furthermore, Margić (2017:51) reported that 'open-mindedness' correlated with positive views on non-native accents, with the participants who were more aware of English language pronunciation diversity being more likely to accommodate and take equal responsibility for any communication breakdown with non-native speakers, as they would with a fellow native speaker. If this familiarity with alternative accents does not exist, it may explain

why deviations from a 'standard' model are viewed negatively, and therefore as harder to process. These findings imply that the background of the listener is often critical in a person's evaluation of a speakers' pronunciation, rather than any particular pronunciation feature.

However, it is not the case that only native speakers view non-native pronunciation negatively. It has been observed by Major et al (2005:44) that non-native speakers may also have a bias against non-native pronunciation of English - whether or not they speak with a native-like pronunciation themselves.

2.3.2 Non-native speaker attitudes towards non-native speech

Derwing (2003) investigated how 100 non-native English speakers that recently migrated to a predominantly monolingual English area of Canada evaluated non-native English pronunciation based on their own experiences as second language users. 55 participants felt pronunciation played a large role in their communication difficulties, but when asked to define precisely what kind of pronunciation difficulties they encountered, 39 were unable to identify anything specific. Out of the participants who did identify a difficulty, 79% identified one or two sounds (such as the production of /θ/ or distinguishing between /l/ and /t/), which were unlikely to have an impact on overall intelligibility. Regardless of whether participants reported that they had pronunciation difficulties themselves or not, 97% believed it is 'important to pronounce English well,' and 95% reported that they aimed to sound like a native speaker. These results cannot be applied to all second language speakers as this particular group had recently moved to an English-speaking culture that they wanted to integrate into, and displaying a pronunciation difference within their own spoken language may have distinguished them as outsiders from their communities of practice. However, these results show that even when pronunciation difficulties are reported, it is difficult to identify exactly what is causing the difficulty, and a known deviation from an expected 'standard' pronunciation is the easiest way to define it. Therefore, the differences in pronunciation and the questionable proficiency of the non-native speaker are perceived as the cause of any breakdown, rather than simple misinterpretation, which frequently occurs in native to native speaker talk (Smith 2009:24).

However as noted in Hendriks et al (2017), often within accent evaluation studies there is little mention of the degree of accent the speakers evaluated have, and how a stronger or weaker accent may result in a more positive or negative evaluation. This may explain why Derwing (2003) was unable to find a consistent feature that was most impactful in determining a negatively or positively evaluated pronunciation. Looking at degree of accentedness specifically, McKenzie (2008) found similar findings to previous research among Japanese participants who evaluated the pronunciation of 2 UK, 2 US and 2 Japanese (1 strongly and 1 mildly accented) speakers, who were more accepting of the Japanese speakers only if they were familiar with the region of Japan each speaker came from, regardless of the degree of accent. Otherwise, there was a tendency to prefer the native English varieties. From these results, McKenzie (2008:81) suggested that giving learners a high level of exposure to a combination of English varieties, both native and non-native, could lead to 'greater acceptance' of diverse international varieties of Englishes. Based on the literature reviewed so far,

this suggestion of increasing the sociolinguistic competencies of second language learners may offer a solution to the ingrained association of RP and GA being what the ideal English speaker should sound like.

However, a limitation of both Derwing's (2003) and McKenzie's (2008) research is that they tackle the issue of non-native evaluations of Englishes by speakers who could be considered as second language learners, rather than second language users. McKenzie's (2008) intention was to find out whether Japanese English pronunciation is evaluated as acceptable for a classroom environment, and Derwing (2003) examined migrants who were attempting to adapt to their new home and culture. Neither of these studies examined highly proficient speakers of second language English, who are unlikely to ever have the intelligibility of their own speech questioned.

Within this section, past research on the evaluation of pronunciation has made conclusions based on qualities of the listener, rather than singling out what features appear to determine 'good,' or 'correct,' pronunciation. Furthermore, the degree of the accent of speakers has often been inconsistent. This study addresses a gap within current literature by evaluating the features that second language users consider as defining 'good' English pronunciation. Using mildly accented speakers with high intelligibility means that the comprehensibility of the speaker is less likely to be questioned, and there will be greater focus on the features that most impact participants' evaluations. Focusing on Dutch second language English users gives this study the unique position of examining the views of people who are already highly proficient English users, and have learnt English as a lingua franca in a country with extremely high standards of proficiency. Furthermore, this particular group of future English teachers will give an insight not only into the current situation of second language evaluations of English pronunciation, but also what the future may hold.

3 Pronunciation Evaluation Task

3.1 Method

The pronunciation evaluation task collected both quantitative and qualitative data with the purpose of observing how students of English evaluated the ‘correctness’ of pronunciations of four distinct global Englishes, and identifying which particular pronunciation features made these evaluations positive or negative.

3.1.1 Participants

30 participants took part in the pronunciation evaluation task: 23 females and 7 males. All participants reported Dutch as their first language and English as their second language. 15 participants also reported speaking at least one other second language to at least a conversational level. 27 participants reported living the Netherlands for their entire life, and the remaining 3 reported they had lived in the Netherlands for more than 5 years. No participants had recently moved to the Netherlands, meaning that all had experienced a similar Dutch secondary school education. The average age of the participants was 22.2 (SD 2.39). A summary of participants can be found in Fig. 3.1.

Figure 3.1: Summary of participants for the pronunciation evaluation task.

Number of participants	Age range	First language (n)	How long living in the Netherlands? (n)	Reported another L2 as well as English
30	18-27	Dutch (30)	Entire life (27) More than 5 years (3)	15

The pronunciation evaluation task was distributed to first, second and third year BA students of English Language & Culture at Universiteit Leiden, all of whom had some background in studying English pronunciation from the first year of their course.

3.1.2 Stimuli

Participants were asked to respond to four short audio clips recorded by four different English speakers, all of whom were of different nationalities and had distinct pronunciations of English. All audio clips were made in a recording booth using Adobe Audition at 44.1kHz.

The Four Speakers

The recorded speakers were selected based on a three-levelled criterion that was influenced by Collins et al's (2011) hierarchy of transfer errors in spoken language. The aim was for all speakers to be highly intelligible, have low irritation or distraction transfer errors, but have a noticeable accent. If speakers matched all three levels of criteria, they were concluded to have a 'mild' accent. A mild accent was preferred to ensure that the speakers represented a noticeable variety of English accents, yet the strength of the accent did not impede understanding. These criteria are explained in Fig. 3.2.

Figure 3.2: An explanation of the criteria used to ensure the speaker recordings were suitable for the pronunciation evaluation task.

Collins et al's (2011) Hierarchy of Error	Level 1: Pronunciation errors cause breakdown in intelligibility	Level 2: Distortion causes distraction or irritation	Level 3: Errors are detectable but do not cause irritation
Criteria for each speaker	High Intelligibility	Low distraction or irritation	Detectable accent

The speakers were checked against these criteria in two stages. Firstly, I used my intuition as a first language speaker of English to decide which speakers should be recorded. Secondly, the subsequent recordings were sent to 3 first language and 3 second language English speakers who were asked to judge whether each speaker had met the criteria set. All speakers were concluded to meet the criteria, and have a noticeable but mild accent of English.

Two first language and two second language speakers of English were chosen to produce the recordings. The speakers shared a similar educational background and were of a similar age (22-27). All were highly proficient in English, as they were currently completing Master's degrees taught in English at Universiteit Leiden. Each speaker had been taught norms of British English (or RP) as opposed to American English as either a first or second language. The decision to only include British influenced pronunciations was made to ensure consistency between the taught styles in pronunciation of each speaker. Inclusion of American influenced pronunciations would have required a larger sample of speakers to ensure a fair representation of pronunciations from each 'standard English language' influence.

Throughout this study, the four speakers will be referred to by their respective cities of birth rather than their nationality, to avoid any generalisations about the variations of English within those countries. The first language speakers were a male volunteer from London (monolingual speaker) and a female volunteer from New Delhi (bilingual in Hindi and English). These two participants were chosen to represent 'inner circle' and 'outer circle' English pronunciations, and two types of first language international Englishes that have higher and lower statuses internationally (Kachru 2009). The second language speakers were a male volunteer from Groningen (first language Dutch) and a female volunteer from Zhengzhou (first language Mandarin). These two participants were chosen to

represent second language users of English in the ‘expanding circle’ that had either a similar educational background to the participants within western European culture (Groningen) and a non-European educational background (Zhengzhou). All speakers identified with English being either their first or second language, as they have been labelled in this study. A summary of each speaker can be found in Fig. 3.3.

Figure 3.3: Summary of the four speakers recorded for the pronunciation evaluation task.

Speaker	Age	Gender	Identifies as L1 or L2 English speaker?	Inner, Outer or Expanding Circle	Other languages
London (UK)	22	Male	L1	Inner	-
New Delhi (India)	27	Female	L1	Outer	Hindi (bilingual L1)
Groningen (The Netherlands)	24	Male	L2	Expanding	Dutch (L1)
Zhengzhou (China)	23	Female	L2	Expanding	Mandarin (L1)

The Recording Material

The recordings consisted of four short literary excerpts from the book *Lost Horizon* (Hilton. 1933). During the recording sessions, each speaker was instructed to read all excerpts aloud twice, as naturally as possible. Before and after reading these excerpts, speakers were presented with instructions to give directions between two landmarks in Leiden, to help them become comfortable with the recording environment. These distractors were not included in the final recording. A full list of the 4 excerpts used can be found in Fig. 3.4.

Figure 3.4: List of excerpts from *Lost Horizon* (1933) used in the recordings.

Page	Excerpt from <i>Lost Horizon</i>
92	<i>One thing he decided instantly; the cold thrill of discovery must not yet be communicated - neither to his companions, who could not help him, nor to his hosts, who doubtless would not.</i>
15	<i>It was a calm night, starry and very warm, and the sea had a pale, sticky look, like condensed milk.</i>
17	<i>Tomorrow, you may actually find it even more interesting. And as for rest, if you are fatigued, there are not many better places in the world.</i>
147	<i>There came a time, he realised, when the strangeness of everything made it increasingly difficult to realise the strangeness of anything; when we took things for granted merely because astonishment would have been as tedious for ourselves as for others.</i>

These excerpts were chosen because the length of the sentences ensures that the prosodic features are noticeable. Literary text was chosen so that the content would not represent a particular speaker identity, but a story. The content was identical for all speakers so that participants would focus on individual pronunciations of each speaker rather than the quality of the content.

Certain features were included in the excerpts in an attempt to prompt non-standard pronunciations of English. Only common Dutch transfer errors of English were purposely prompted as, theoretically, these would be the problematic areas that the participants of the pronunciation evaluation task would have personal experience of. The Dutch transfer errors included in these excerpts were chosen from Collins et al's (1987:93-96) Accepted English Pronunciation, in the section entitled 'problems of Dutch speaking learners of English.' A total of 24 common segmental deviations from RP pronunciation were marked as being 'crucial' or 'serious errors' for Dutch learners of English. Each of the recordings was analysed for the above transfer errors. Any deviations found in the recordings are reported in Fig. 3.5.

Figure 3.5: Crucial or serious 'common Dutch errors' from Collins et al (1987) found in the recordings.

Sound	Description of transfer error	Speakers that made at least one 'error'	Example from recording	Page of excerpt
t	Lack of final glottal reinforcement	Groningen	Not - /nɒd/	17
ð	Replaced by /d/ in initial positions	New Delhi Groningen	There came - /dɛər keɪm/	147
	Replaced by /d, z, t, s/	Zhengzhou	Everything - /evɔ:ri:stɪŋ/	147
w	Replaced by /v/	New Delhi Groningen	Very warm - /veri: vɔ:rm/	15
æ	Replaced by Dutch vowel /ɛ/	Groningen	Starry - /stɛ:ri:/	15
l, ɫ	/l/ used instead of /ɫ/	Zhengzhou	Thrill - /θrɪl/	92
aɪ	Elongation of the vowel /aɪ/	Groningen	Night - /naɪ:t/	15
ɪ	Too close in final position (like Dutch /i/ in <i>koffie</i>)	Groningen	Instantly - /ɪnstəntli/	92

Although other deviations from RP are to be expected in all four speakers, they were not attempted to be prompted, but left to occur naturally. The participants are taught to avoid common Dutch

transfer errors, and these transfer errors were purposely included to ensure there were enough possibilities for them to occur in each of the recordings. This would allow the investigation to note whether participants focused more on common Dutch transfer errors regardless of which speaker they evaluate. Any other pronunciation deviations that would be seen as ‘incorrect’ would be entirely the perception of the participants. Moreover, a short phonetic analysis of the recordings was conducted to observe what sounds were most likely to be perceived as deviant, based on the participants’ knowledge of English. This was completed through a comparison of the recordings to Jenkin’s (2007:23-25) description of pronunciation variation within global Englishes, and included any additional comments on any striking features of the recordings. This information is represented in Fig.3.6. Any common Dutch errors that were referred to in Fig.3.5 were not included to avoid duplication of information.

Figure 3.6: Short phonetic analysis of the 4 recordings showing the features most likely perceived as deviant that are not common Dutch errors, through a comparison with Jenkins’ (2007) descriptions of pronunciation variation in global Englishes

Speaker	Variation described in Jenkins (2007)	Other comments
London	Word-final glottal stop replacing /t/	-
New Delhi	Occasionally little difference between short and long vowels	-
Groningen	-	Final rising intonation pattern
Zhengzhou	Diphthongs occasionally pronounced as monophthongs	Segmented intonation Occasional addition of /h/ in the onset position

The final stimuli consisted of four audio clips lasting approximately 40 seconds each. These were edited versions of the full recordings, in which the readings with the least amount of hesitation were chosen. All recordings were edited using *Praat* (Boersma & Weenink 2018). Information on how to access the recordings can be found in appendix 1.

3.1.3 Procedure

The pronunciation evaluation task was created as an online survey using Qualtrics (2018). The survey was conducted through an anonymous URL that was distributed to participants via an email invite from their English instructors. An introductory page stated what was expected of participants, and how they would complete the survey. Participants were informed that all responses would remain anonymous. All instructions were given in English.

Background Questions

Participants were first asked to complete a list of demographics questions, including reporting their first and second languages. This was to ensure that the linguistic background or potential multilingual influences of the participants could be identified.

Evaluation Tasks

The survey was designed so that participants would encounter the recordings in a randomised order to ensure that ordering effects would not impact the results. Each participant was instructed to listen to the entire recording, and then to answer questions about the pronunciation of the speakers. All questions were identical for each speaker. (Please see appendix 1 for a link to the survey).

Qualitative data was collected through open-ended questions regarding the participant's evaluation of each speaker's pronunciation. Quantitative data was collected through Likert scales of 1 to 8, in which participants were asked to rate four different segmental and prosodic features of the speaker's pronunciation. The scale of 1 to 8 was chosen so that participants would not associate the scale with the grading system of the Netherlands (in which all scores are between 1 and 10, with anything over 6 being deemed a pass), as that may have influenced participants to avoid giving low scores due to the association of scores lower than 6 being a fail. This choice of scale also ensured there would be no central number and avoided neutral answers; participants needed to make a decision between either a low score (1-4), or a high score (5-8).

The pronunciation features mentioned in the Likert scales were: *pronunciation of individual words*, *intonation*, *use of correct consonants*, *use of correct vowels*, and *accent consistency*. These five categories were influenced by Hoorn et al's (2014:105) study, investigating how teachers in the Netherlands grade the pronunciation of English students. Their results included the most commonly quoted reasons for giving a student a high or low score, and the top 5 were taken as a benchmark for the Likert scales in the present study.

A second set of Likert scales were used as participants were also asked to judge to what degree they could understand the speaker, and to what degree they thought the recording was of a native speaker. The same scales of 1 to 8 were used. Although all speakers were categorised as being highly proficient in English (see section 3.1.2) and low scores are not expected in the intelligibility question, this question was asked to compare whether low scores in the 'correctness' of a speaker's pronunciation would correlate with the perceived comprehensibility of their speech.

3.2 Results

The following results are presented in the same order as the questions in the pronunciation evaluation task. The results for all speakers will be compared for each question, with a focus on the themes that arose in the responses.

Individual participants will be identified by number, chronologically. For example, the first participant that took part in the evaluation task will be referred to as P1, and the final participant P30.

3.2.1 The overall classification of speakers

In the first question of the evaluation task, participants were asked to make a judgement about whether the recording they heard was 'good' or 'bad'. The responses were divided into positive, neutral or negative categories. Positive responses included classifications such as: perfect, very good, quite good, everything is clear. Negative responses included classifications such as: bad, not good, mediocre, not natural. Neutral responses included the cases in which participants did not give a clear positive or negative view, including responses such as: ok, understandable. 3 responses were discounted due to their ambiguity, as it could not be determined whether the comments were positive, negative or neutral. For example, P27 responded to this question for the New Delhi speaker by saying 'it is very clearly non-native English.' These results are shown in Fig. 3.7.

Fig. 3.7: Comparison of overall positive, negative and neutral reactions to each speaker. Overall participants who gave a response for each speaker were London:30, New Delhi: 27, Groningen:29, Zhengzhou:29. The blue box marks the highest number of responses for that speaker.

Score	London	New Delhi	Groningen	Zhengzhou
Positive	30	23	15	3
Neutral	-	2	7	6
Negative	-	2	7	20

These initial judgements indicate a preference for the pronunciation of the first language English speakers, London and New Delhi, with 100% and 85% of positive responses respectively, over the second language English speakers, Groningen and Zhengzhou, who received 50% or less of positive responses.

It is also notable that the Groningen speaker has a clearly preferred pronunciation to the Zhengzhou speaker, with 40% more positive responses. However, due to the wide distribution of answers for the Groningen speaker, with 50% of participants classifying the speaker as either positive or not positive, there is no clear trend in how this speaker was viewed.

The next section will look into this proposed hierarchy of preferred pronunciations in more detail.

3.2.2 Ranking features of pronunciation

Participants were asked to rank the following pronunciation features on a Likert scale of 1 to 8 (with 1 being very low ability, and 8 very high ability): *pronunciation of individual words*, *intonation*, *use of correct vowels*, *use of correct consonants*, and *accent consistency*. In this section, a summary of the main statistical findings will be presented. The outcome of each test will be fully evidenced in the appendix.

General trends for the ranking of pronunciation features on Likert scales

As ordinal data was collected, non-parametric tests were used to test the statistical differences between the four speakers and five scales. To firstly gain an indication of the general trends in the data, the median scores of each speaker for each scale were plotted in Fig. 3.8.

Fig. 3.8: Comparison of median scores rating pronunciation features for each speaker.

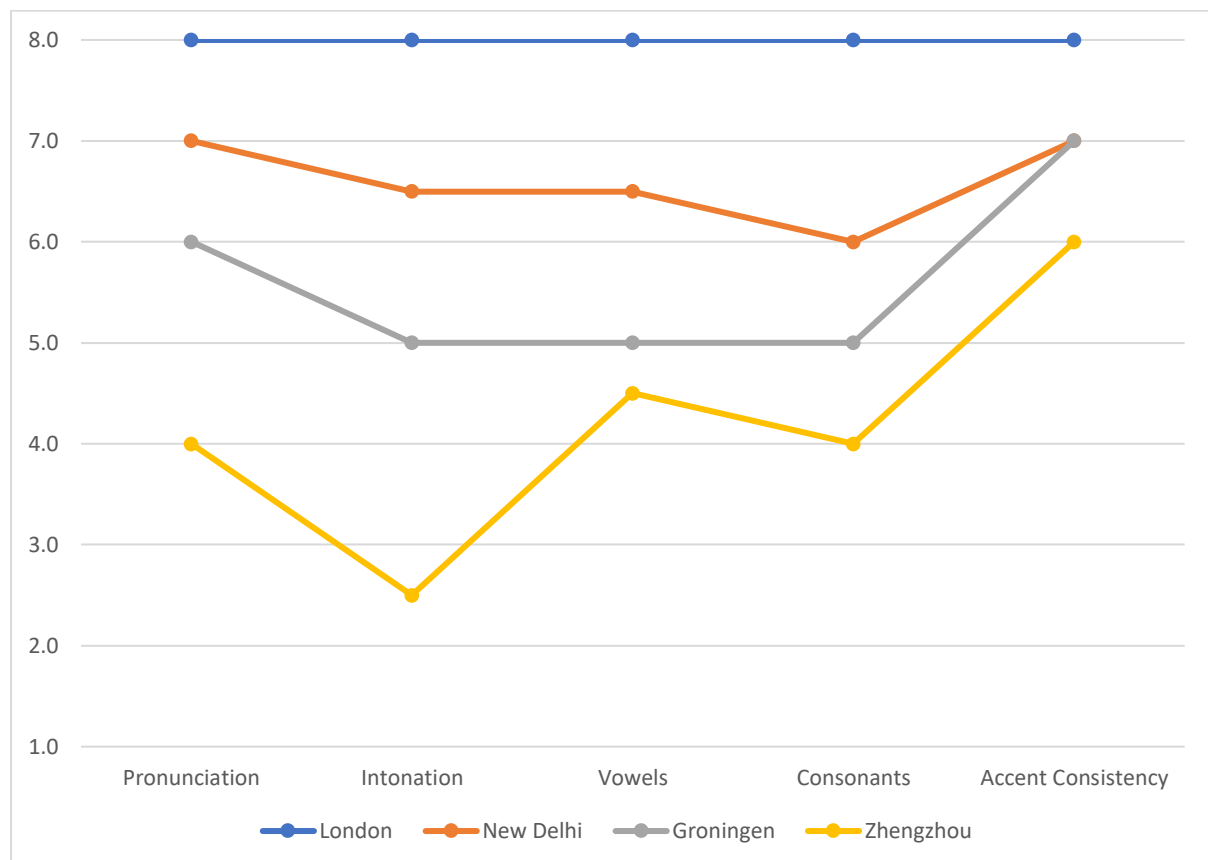


Fig. 3.8 suggests that there is no particular feature which appears to be a problem for pronunciation

(resulting in lower scores) that is consistent for all speakers. The median scores replicate the hierarchy of preferred pronunciation which was displayed in Fig. 3.7, with the London speaker being evaluated most positively, followed by the New Delhi, Groningen and Zhengzhou speaker.

Five Friedman's ANOVAs were then used to compare results between speakers for each individual scale, such as *pronunciation of individual words*. A significant difference was found for each of the five scales, with $\chi^2(3) = 64.6$, $p < .001$ for pronunciation of individual words, $\chi^2(3) = 52.1$, $p < .001$ for intonation, $\chi^2(3) = 52.5$, $p < .001$ for correct use of vowels, $\chi^2(3) = 49.9$, $p < .001$ for correct use of consonants, and $\chi^2(3) = 21.2$, $p < .001$ for accent consistency. Wilcoxon signed rank tests were used to follow up these findings, and a Bonferroni correction was applied to account for all 6 comparisons that needed to be made, meaning that all effects are reported at a .008 level of significance. The main findings of the post-hoc tests will be summarised below. (Please see appendix 2, table 1 for the full statistical results).

It was found that for the *pronunciation of individual words*, *intonation*, *vowels* and *consonants* scales, almost all scores between speakers were significantly different. Only the Groningen and New Delhi speaker were not significantly different from each other for these scales, which was thought to occur as both of these speakers have distributions that are neither consistently high (such as the London speaker), nor consistently low (such as the Zhengzhou speaker). Other than this one overlap, there are significant differences for almost all speakers in almost all scales, meaning that the difference in scores for each category seems to be mostly predictable based on which speaker is being judged, rather than the category they are being judged on.

On the other hand, in the *accent consistency* scale, the Zhengzhou speaker was the only speaker that was significantly lower than the London speaker ($Z = 3.50$, $p < .001$), as well as New Delhi speaker ($Z = 2.72$, $p = .006$). No other significant difference was found between the speakers. In this case, the little significant difference between all speakers may be expected as the *accent consistency* scale is not asking participants to evaluate a particular trait of pronunciation positively or negatively; only whether the speaker uses their accent consistently. Even in the two cases the Zhengzhou speaker is significantly different, as there is no significant difference between the Zhengzhou and Groningen speaker, or the Groningen speaker and any other speaker, we may conclude that the difference in *accent consistency* scores was overall quite low. This finding suggests that the accents remained consistent throughout the recordings, and any deviations in pronunciation were systematic for that speaker.

General trends for the individual scores of speakers

As there was a consistent difference between almost all speaker's scores for each scale, to further test whether any particular aspect of pronunciation is felt to determine the overall attitude towards a speaker, four Friedman's ANOVAs were used to compare the scores between each speaker for the individual scales. No significant difference was found between the scores of any of the scales for the London speaker, with $\chi^2(4) = 4.46$, $p > .05$. A significant overall difference was found for the New Delhi, $\chi^2(4) = 26.17$, $p < .001$, Groningen $\chi^2(4) = 34.32$, $p < .001$, and Zhengzhou speakers, $\chi^2(4) = 45.19$, $p < .001$. Wilcoxon signed rank tests were used to follow up these findings. A Bonferroni

correction was applied to account for all 10 comparisons between scales that needed to be made, meaning that all effects are reported at a .005 level of significance. The main findings of the post-hoc tests will be summarised below. (Please see appendix 2, table 2 for the full statistical results).

For all three speakers, *accent consistency* was significantly higher than all speaker's scores for *intonation*, *use of correct vowels* and *use of correct consonants*. As discussed above, this category is a rating that does not reflect either a positive or negative evaluation of 'correctness' of the speakers' pronunciation, and may therefore be expected to be inconsistent with the scores from other scales. The reason this was not significantly different for the London speaker, is because this speaker had consistently high scores for all scales, as can be seen in Fig. 3.8.

Intonation was found to be significantly lower than all other scale scores for the Zhengzhou speakers, and significantly lower than *overall pronunciation* for the Groningen speaker. This finding may suggest that *intonation* is an important feature for pronunciation as it is the only scale which is significantly lower than other scales for the two non-positively rated speakers (Groningen and Zhengzhou). The only other scale which was found to be significantly different was that *pronunciation of individual words* was a significantly higher score than *use of correct consonants* and *use of correct vowels* for the Groningen speaker. Excluding the *accent consistency* scale, there are only 6 cases (out of a total of 18 comparisons) of significantly different scores between all scales for individual speakers. 4 of these 6 cases are the *intonation* scale being significantly lower than the other scales for the two second language speakers. These results provide evidence, again, that the scores are mostly predicted based on which speaker is being evaluated, and not which scale they are being judged on.

The next few sections examine the qualitative data collected, to present an in-depth look into what particular pronunciation features were examined by participants.

3.2.3 Determining which features impact pronunciation the most

This section will compare the elements of pronunciation that were categorised as 'features to improve' for all speakers; observing any themes that arise within these critical comments. Afterwards, I will compare the elements of pronunciation that were categorised as 'good' features of pronunciation for all speakers. Any other themes arising from the comments about speakers' pronunciations will be discussed in the following sections.

All responses were coded for the feature they represented, in the categories that were most commonly mentioned: *segmental deviation*, *intonation* and *fluency* (including rhythm and voice quality). Any additional features that did not fit into these categories were marked as *other*. As over half of the segmental deviations from expected sounds were reported as errors, this category was further divided into *pronunciation of consonants* and *pronunciation of vowels*. If participants reported a particular consonant to improve (rather than generally commenting that the speaker could work on their consonants), these phonemes were also collected to observe whether any particular phoneme was particularly problematic across multiple speakers. The results for the features to improve are plotted in Fig. 3.9.

Fig. 3.9: Comparison of the frequency of features to improve for all speakers. Comments were coded for the most commonly mentioned features.

NEGATIVE COMMENTS					
Speaker	Consonants	Vowels	Intonation	Fluency	Other
London	5% overall which includes: 3% /ʀ/	5% overall which includes: 4% 'non-standard'	2%	8%	5% accent consistency
New Delhi	47% overall which includes: 13% /θ/ 6% /t, l/ 6% aspiration 5% /v, w/	7% overall	2%	8%	3% word stress
Groningen	39% overall which includes: 12% devoicing 7% /θ, ð/ 6% /t, s, f/	20% overall which includes: 4% /æ/ 4% /ɒ/	17%	8%	4% influence from Dutch
Zhengzhou	37% overall which includes: 8% /θ/ 7% /v/ 6% /r, t/	13% overall which includes 2% /i:, æ/	23%	12%	6% incorrect word boundaries 5% word stress 3% comprehensibility

The two categories that appear to divide the first and second language speakers for features to improve are *intonation* and *pronunciation of vowels*. In both of these cases, the London and New Delhi speakers have very low mentions of improvements to be made, whereas the Groningen and Zhengzhou speakers both have high percentages of improvements to made.

Despite this, *pronunciation of consonants* had a high percentage of comments to be a point of improvement for all speakers except the London speaker. 47% of comments regarding the New Delhi speaker's pronunciation noted the deviation in expected consonant sounds as speaker errors. An explanation for such a high percentage may be that few other pronunciation errors were observed for the New Delhi speaker, so all negative comments were concentrated in this category. The individual phonemes reported as errors generally did not overlap between speakers, however the /θ/ was reported as an error in 7% or more of cases. Jenkins (2007) stated that /θ/ is one of the most difficult phonemes to replicate in English, but that a deviation in pronunciation has little impact on intelligibility. Furthermore, De France & Smakman (2013:8) noted certain deviations in consonants are highly stigmatised, resulting in negative evaluations even if the specific consonantal

pronunciation does not affect communication. This recognisable difficulty of /θ/ may mean the sound is marked for these second language participants as being associated strongly as a sound of English, and any deviation may be observed as a sign of a speaker not being English.

The Groningen speaker had specific comments regarding interference from their first language (L1) Dutch, which were often generalised as ‘the pronunciation is too Dutch-like.’ Furthermore, it is also observed that specific Dutch transfer errors were commented on as obstruent-final devoicing had a large amount of mentions (12%) as segmental deviation. This specificity reflects the high recognition of the Dutch speaker, as such specific comments did not appear for the other 3 speakers.

For comparison, the results for the good features of pronunciation are plotted in Fig. 3.10.

Fig. 3.10: Comparison of the frequency of good features of pronunciation for all speakers. Comments were coded for the most commonly mentioned features.

POSITIVE COMMENTS					
Speaker	Consonants	Vowels	Intonation	Fluency	Other
London	25% overall which includes: 5% non-rhotic 3% /θ/	11% overall	13%	15%	6% British accent
New Delhi	5% overall	2% overall	12%	8%	14% comprehensibility
Groningen	0	0	3%	0	3% consistent accent
Zhengzhou	0	0	2%	0	2% comprehensibility

Intonation being the divide between the first and second language speakers is further highlighted in Fig. 3.10. The Groningen and Zhengzhou speakers have a very low percentage of comments that mentioned good intonation, yet the London and New Delhi speakers have a large percentage of comments reporting that their intonation was a sign of good pronunciation. The *pronunciation of vowels* does not appear to have the same divide as suggested in Fig. 3.9, due to the fact the New Delhi speaker also had a low amount of mentions (2%) for this as a positive feature.

The London speaker was highly regarded for their similarity to RP. Within their highest category of good features of pronunciation, *pronunciation of consonants* (25%), the two features that were explicitly commented on were the non-rhoticity of the accent, a defining feature of RP, and the correct use of /θ/. This is further highlighted as within *other*, 6% of comments specifically mentioned that they thought the speaker had a ‘British’ accent. This is interesting to compare to the Groningen speaker in Fig. 3.9, whose Dutch accent was seen as a feature to improve.

Comprehensibility was only mentioned as a pronunciation feature for the non-European speakers. It

is interesting to note that the New Delhi speaker had a high percentage of comments about their comprehensibility (14%), which is not necessarily a specific feature of the pronunciation. This suggests that despite the fact consonant deviations were highly remarked as features to improve, the ease to comprehend the speaker may explain why the New Delhi overall had the second most preferred pronunciation (see sections 3.2.1 and 3.2.2).

Additional comments about each speakers' pronunciation were also assessed, and the following themes were found which will be discussed in the below subsections.

3.2.4 Identification of the pronunciation

For both the London and Groningen speakers, most participants attempted to identify the speaker's nationality. The most recognisable pronunciation was that of the Groningen speaker, who was labelled as being Dutch by 66% of participants, with a further 13% of participants labelling them as being a non-native English speaker. Although the London speaker was never correctly labelled as being from London or even more generally from the south of England, 40% of participants labelled this speaker as being from either the UK or Ireland. On the other hand, participants only attempted to label the New Delhi and Zhengzhou speakers 6 times, combined. In particular, within the 3 attempts to label the Zhengzhou speaker, only 1 correctly identified them as being Chinese, with the other 2 using the general term of 'Asian' - a concept which would be extremely hard to define. This result suggests the latter two pronunciations were less familiar for participants than the former.

There was a contrast in how speakers were evaluated depending on whether they were labelled or not. This is especially evident with the London speaker. Participants who labelled the London speaker as being 'British' explicitly stated that the pronunciation was as they expected, and it was often the case that 'British sounding pronunciation' - which in this case is assumed to refer to RP - was used as a positive description of this speaker. However, in the cases in which the London speaker was identified as not being British, participants provided suggestions of how the speaker could improve to be more RP-like. This is evidenced in Fig. 3.11.

Fig 3.11: The contrast in comments after identifying the London speaker as native or non-native

Evaluation	Participant	Comments	Overall average on Likert scales (maximum 8)
Positive	P13	<i>It's good, it's really British.</i>	8
	P18	<i>You can hear that the speaker probably is a native speaker from the United Kingdom. Therefore, the pronunciation is very good.</i>	8
Negative	P3	<i>If the speaker's goal is to sound more native then I would suggest trying to practice the British accent more.</i>	7.2
	P21	<i>I think the speaker has mastered the British accent quite well but is not a native speaker. I think there are some slip ups that give them away (intonation, inconsistency).</i>	5.6

The high recognition of the Groningen speaker resulted in comments that addressed his need to sound less Dutch in order to have a better English pronunciation. This suggests that participants saw obvious interference from Dutch on to English as a large error, and as they have the most experience with these two languages, were able to expect and point out specific transfer errors more easily. Examples can be found below.

P20 - *His final consonants are devoiced, like in 'condensed'*

P26 - *Distinction between E and A is not clear. The A sounds like Dutch E, which is not the case in English.*

P2 - *The pauses between the words do not sound natural: try to imitate the rhythm with which English speakers talk.*

In comparison, the two unfamiliar pronunciations of the New Delhi and Zhengzhou speakers meant that in most cases both speakers were labelled as being 'foreign,' and therefore as second language speakers of English. This theme was only absent when the New Delhi speaker was correctly identified as being either 'Indian' or more generally as 'a former colony of the UK,' and in these cases participants referred to the fact this speaker was therefore a first language speaker of English. However, the New Delhi speaker was only identified by 3 participants. Evidence of the labelling of 'foreign' for both speakers can be found in Fig. 3.12.

Fig. 3.12: Table showing comments which labelled the New Delhi and Zhengzhou speakers as 'foreign'

Speaker	Participant	Comments
New Delhi	P29	<i>It's good ... You can hear that it is not a native speaker though.</i>
	P19	<i>Good but definitely foreign.</i>
	P4	<i>Good but improvements can be made.</i>
	P18	<i>The pronunciation is good enough for me to understand the words, but I did have to listen more carefully than I normally would. It needs some improvement, but I would not consider it to be bad.</i>
Zhengzhou	P28	<i>She speaks English with an Asian tone, which does not go together.</i>
	P11	<i>The slow speech and the unusual pausing patterns make the speech clearly sound non-native.</i>
	P27	<i>It is very clearly non-native English.</i>

Fig. 3.12 shows that the New Delhi speaker's pronunciation is regarded positively whereas the Zhengzhou speaker's is not. Despite this, the comments show that both speakers are labelled as being 'foreign,' and even for the New Delhi speaker, who is a first language English speaker, this perception of 'foreignness' is seen as an area to improve.

3.2.5 Association of RP as being English

As discussed within section 3.1.2, this study chose to focus on the English pronunciations of first language speakers with a connection to the UK and second language speakers who had been taught 'British English' (RP). However, there was no specific indication within the pronunciation evaluation task that all speakers had some connection to RP or the UK. This decision was made to ensure all participants were free to make conclusions or comparisons of each speaker based on their personal views on English pronunciation.

Despite this, deviations from RP consonants and vowels in the 4 recordings were most often reported as being errors. GA was occasionally mentioned as an acceptable form of English, but this was only ever suggested as a side comment by a few participants. Despite the fact RP is a minority accent in the UK (Szpyra-Kozłowska. 2015:31), participants often referred to RP pronunciation more generally as a British accent, and felt that any segmental deviations were features to improve and often limited their negative evaluations to only describe this type of transfer error. Any segmental errors that were suggested as areas to improve often referred to specific sounds. There was no suggestion that these segments caused unintelligibility for any speakers except the Zhengzhou speaker, and as the participant recognised what was meant to be spoken in each case, it is highly unlikely any actual unintelligibility was caused by segmental deviations. Examples of these suggestions for all speakers can be found in Fig. 3.13.

Fig. 3.13: Examples of suggested pronunciation improvements for all speakers

Speaker	Participant	Comments
London	P23	<i>Some of the vowels (such as in "warm" or "because") seem a bit short.</i>
	P11	<i>More emphasis on the consonants, such as 'T' in the word 'NOT'</i>
New Delhi	P19	<i>Pronunciation of th is incorrect.</i>
	P6	<i>The pronunciation of some letters, especially the letter r, was very clearly not English.</i>
Groningen	P3	<i>Try to work more on you aspiration on dental vowels (t,d etc.) They sound very Dutch like. This small adjustment will make you sound more native.</i> <i>Improve the pronunciation of vowels (sounds very Dutch, e.g. "difficult" sounds like "difficol't").</i>
	P7	
Zhengzhou	P15	<i>The /s/ sounds non-native.</i>
	P23	<i>The speaker needs to pronounce "th" correctly (which she actually seems to do in "anything" at some point)</i>

There were, however, a few small exceptions to this theme that only appeared in the comments for the New Delhi speaker. As reported above, the New Delhi speaker was most often regarded as being 'foreign' and therefore usually as a second language speaker of English. However, some participants took back their earlier comments that the errors should be improved, due to the speaker's labelled 'foreignness' and high intelligibility of English. These responses can be seen below.

P6 - *This is not necessarily something they should improve, since it was not a hindrance to understanding, but it is something that makes it very clear they have a non-British or American accent.*

P27 - *The speaker may well be a native speaker of a non-standard variety of English, in which these differences in intonation and pronunciation are commonly accepted.*

P11 - *The consonants are a bit strange sometimes, but that's part of this accent.*

These comments accepting the differing pronunciation were not present in any form for either the London, Groningen or Zhengzhou speaker, suggesting that the label of 'foreignness' but unquestioned intelligibility of the speaker had made this impact on a few participants. However, even in these few cases all participants still compare the pronunciation to RP or GA, reflecting the high association with these two forms representing the 'correct' pronunciations norms of English, even when participants were accepting of other varieties.

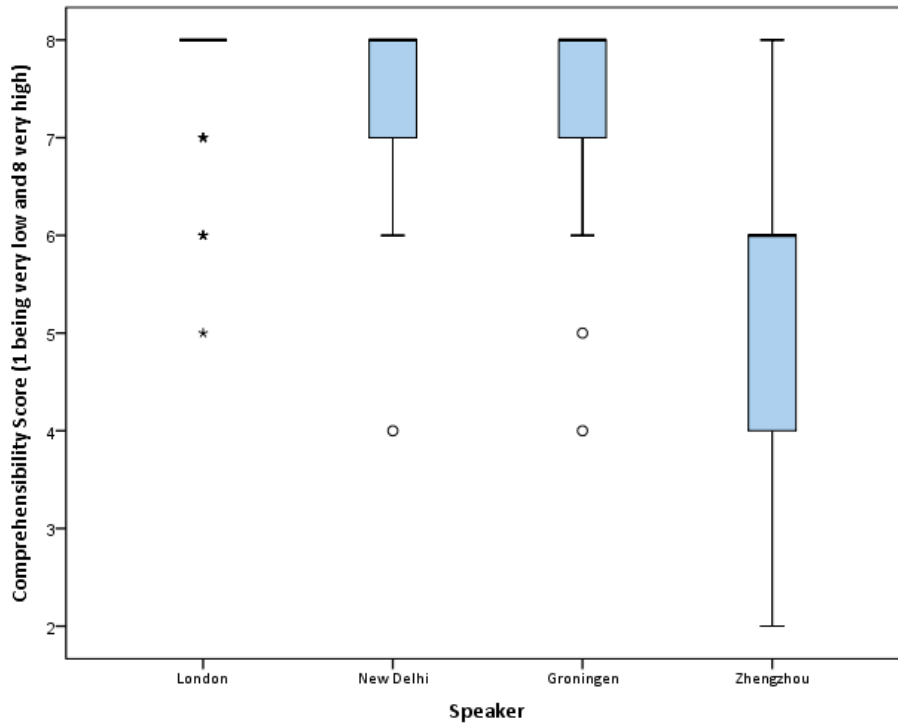
3.2.6 Comprehensibility and native speaker scores

On a scale of 1 - 8 (1 meaning 'not at all' and 8 'absolutely'), the final two questions asked participants to rank how well they understood each speaker, and how much they felt each speaker was of native speaker of English.

Firstly, I will examine the results of the question asking participants to rank how much they understood each speaker.

A Friedman's ANOVA was used to test the statistical difference between the comprehension scores for all speakers. The difference in understanding was found to be significant, $\chi^2(3) = 53.97, p < .001$. Wilcoxon signed rank tests were used to follow up this finding, and a Bonferroni correction was applied to account for all comparisons, meaning all effects are reported at a .008 level of significance. The Zhengzhou speaker was significantly less understood than all other speakers, at $Z = 4.6, p < .001$ in all cases. There was no significant difference between other speakers' scores. This finding shows that there is little difference between the understanding of the London, New Delhi and Groningen speakers, all of whom were rated highly, and only the Zhengzhou speaker was significantly less well understood. These findings were plotted in Fig. 3.14 for further examination.

Fig. 3.14: Results from the question: 'How well do you understand this speaker?'
(1 = not at all, 8 = absolutely)



In Fig. 3.14, the average distribution of the scores is shown by the blue plots, with the whiskers showing the remainder of the distribution. The London speaker has a thick line marking the majority of scores, as there was little distribution in their results. The thick lines at the top of each box mark the median point of scores. Any outliers are marked with a star or circle.

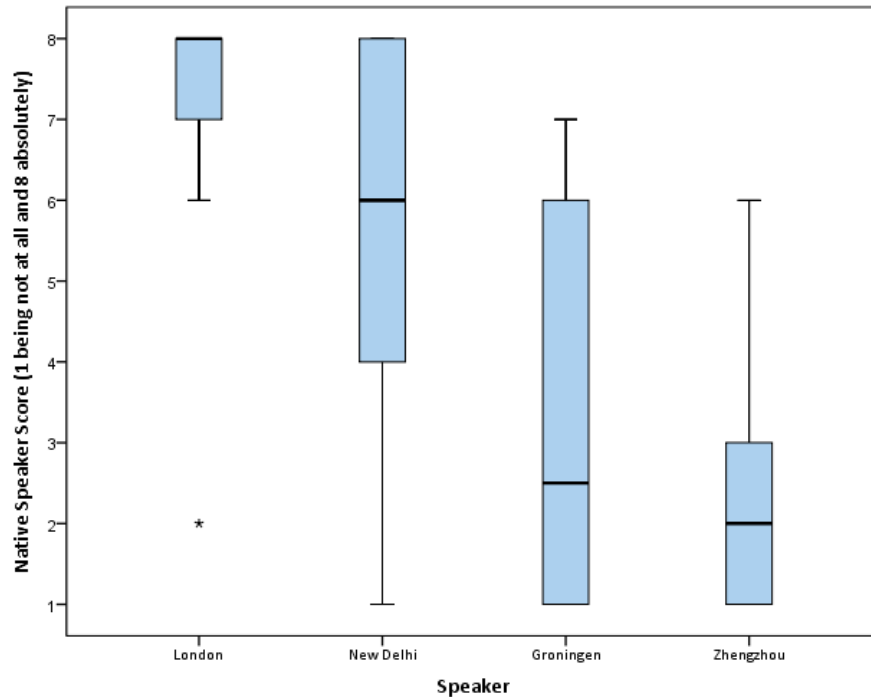
Fig. 3.14 shows a three-way grouping. The London speaker is rated almost exclusively with the highest score of 8. The New Delhi and Groningen speakers have an almost identical distribution of very high scores, with the majority of scores either being 7 or 8. The Zhengzhou speaker is less understood, with the majority of scores being between 6 and 4. There is a large distribution for the Zhengzhou speaker, with some participants marking her the highest score of 8 for understanding, and others marking her the second lowest score of 2. As the results of the Friedman's ANOVA showed that there is no significant difference between the London and New Delhi speaker scores, and the Groningen speaker score are very similar on the boxplot, we can conclude that misunderstanding the pronunciation in the recordings was only a potential problem for the Zhengzhou speaker.

Secondly, I will examine the results of the question asking participants to what degree they thought each speaker was a native speaker of English.

A second Friedman's ANOVA was used to test the statistical difference between the native speaker scores for each of the 4 speakers. The difference in native speaker score was found to be significantly different, $\chi^2(3) = 65.69, p < .001$. Wilcoxon signed rank tests were used to follow up this finding and a Bonferroni correction was applied to account for all comparisons, meaning all effects are reported at a .008 level of significance. The differences between all the native speaker scores

were found to be significant at $p < .001$. (For a full list of these significance levels, please see appendix 2, table 3). This finding shows that whereas the difference between the understanding of the speakers was relatively low (as shown in Fig. 3.5), each speaker was distinctly given their own place on the native speaker scale. These findings are plotted in Fig. 3.15 for further examination.

Fig. 3.15: Results from the question: 'To what degree do you think this speaker is a native speaker of English?' (1 = not at all, 8 = absolutely)



At first sight, Fig. 3.15 shows a two-way grouping. The London and New Delhi speaker have a majority of positive scores, and the Groningen and Zhengzhou speaker have a majority of negative scores. However, I consider that in a similar sense to Fig. 3.14, Fig. 3.15 shows the same three-way grouping. The London speaker is (despite one outlier) regarded to be a native speaker. The Zhengzhou speaker is regarded to be a non-native speaker. In this case, the New Delhi and Groningen speakers do not have a similar plotting or median, however, the similarity between these two speakers is that their scores are highly distributed. Participants appear to be unclear on how to categorise these two pronunciations, with the majority of scores being between 8 and 4 for the New Delhi speaker (with some participants even giving this speaker scores between 4 and 1), and the majority of scores being between 6 and 1 for the Groningen speaker.

These results show the distribution in attitudes towards English speakers who are clearly intelligible but also clearly not either RP or GA - do all native speakers adhere to certain standards, and is clarity of speech enough for a good pronunciation? Earlier results from this experiment suggest the native speaker model is still considered to be of importance, but despite that, these scores show there is still ambiguity in what makes a speaker regarded as being 'native.'

4 Follow-up Interviews

4.1 Method

The follow-up interviews consisted of a small-scale qualitative experiment with the intention of observing how students of English defined 'good' pronunciation without any stimuli.

4.1.1 Participants

10 participants took part in the follow-up interviews. All were female, first language speakers of Dutch, second language speakers of English, and current undergraduate students majoring in English Language at Universiteit Leiden. All participants were aged between 19-24, and 4 participants reported other second languages of Japanese, German and French. A summary of participants can be found in Fig. 4.1 below.

Figure 4.1: Summary of participants for the follow-up interviews.

Number of participants	Age range	First language (n)	How long living in the Netherlands? (n)	Reported another L2 as well as English
10	19-24	Dutch (10)	Entire life (10)	4

4.1.2 Procedure

The participants completed an informal written interview after an English tutorial. Each signed a consent form, which detailed the purpose of the interview, before the interviews took place. Participants were informed that all responses and personal information would be kept anonymous.

The written interviews took place within an open office area, where participants were given a sheet of questions to answer. I stayed in the room with the participants as they completed the interview. They were instructed that this was not a test, and that if they had any questions or would also like to discuss their answers with me orally as well as in writing, they were able to do so. Any oral answers were noted down by me, and added as additional notes for the participant after the interview was complete. The interviews took around 10 minutes to complete.

The interview consisted of 8 open-ended questions which examined how the participants defined good and bad English pronunciation. This included questions about how they viewed their own English pronunciation, as well as how they react to non-native pronunciation of Dutch. A full list of questions asked during the interviews can be found in appendix 3.

4.2 Results

The results will be presented in the order the responses were collected, with a focus on the themes that arose from similar groups of questions.

4.2.1 Being understandable vs. sounding native-like

The first four questions addressed descriptions of ‘good’ and ‘bad’ English pronunciations, as well as asking participants to define one feature they considered the most important for ‘good’ pronunciation, and one feature they considered to be an indication of ‘bad’ pronunciation. Two themes were found within the responses to these questions. The first was that in the first three questions, almost all participants responded that either having an RP/GA pronunciation or being understandable is the most important feature of pronunciation. When ‘being understandable’ was chosen as the most important feature, the response did not define precisely what made a person’s pronunciation ‘understandable.’

Responses to the first three questions were then marked with the following codes: *sounding native-like is important*, *being understandable is important*. Any responses that did not fit into either of these categories was marked as *other*. The results have been plotted in Fig. 4.2.

Fig. 4.2: The frequency of responses about the importance of being understandable or sounding native-like in pronunciation

Question	Being understandable is most important	Sounding native-like is most important	Other
Asked to describe good English pronunciation	4	6	-
Asked about the most important feature of good pronunciation	5	4	1 - (correct grammar)
Asked to name all types of good English pronunciation	1	9	

Participants’ choices showed little consistency in which of the categories they chose for these three questions, illustrating the difficulty in deciding how much of a priority native-like pronunciation is. However, within the last question we can see how strong the association of ‘good’ English as being a native standard is, as only 1 participant referred to a non-native variety of English as being an acceptable English pronunciation - in this case Dutch. Of the 9 participants that mentioned exclusively native varieties, only 1 mentioned a variety that was not either RP or GA, which was Australian English.

The second theme was that ‘bad’ English pronunciation was most often defined as interference from

a first language. Participants were asked to describe a feature that represented incorrect pronunciation in the fourth question. These results were coded using Collins et al (1987) English pronunciation guide for speakers of Dutch, to note whether the features described were specific Dutch errors. The codes used for transfer errors were either: *Dutch specific* or *general transfer*. If a participant did not refer to a transfer error, their response was coded as *other*. The results have been plotted in Fig. 4.3.

Fig. 4.3: Frequency of participants who referred to a general transfer error, a Dutch specific transfer error, or other feature when asked about the most impactful feature for bad pronunciation

Code	General Transfer	Dutch Specific	Other
No of Participants	2	7	1

As Fig. 4.3 shows, only 1 participant did not refer to a transfer error as being the most negative feature of a person’s pronunciation. Instead, they mentioned incomprehensible speech and ‘general mispronunciation of syllable patterns.’ All 9 other participants stated that a transfer error would make an English pronunciation bad, and 7 of these participants specifically referred to a common Dutch transfer error, including the below examples.

P6 - *Dutch speakers tend to rely on their knowledge of Dutch when pronouncing English words which results in bad pronunciation.*

P7 - *Bad pronunciation of English occurs when speakers do not use English phonemes. For example, some people do not pronounce a good /g/ and instead use /x/.*

Although Fig. 4.2 showed that being understandable was perceived as an important part of pronunciation, in Fig. 4.3 deviation from ‘English’ phonemes was seen as an intrusion of the first language, and an error to be avoided. If being understood was indeed the most important factor, the influence Dutch may have on English would not be considered a sign of bad pronunciation if it did not impede understanding. Only P1 referred to potential misunderstanding that may arise from bad speech quality.

The next section discusses the participants’ views on the native model in more detail.

4.2.2 Attitudes on the native speaker model

5 participants reported that they aim to sound like a native speaker of English. Although this is half of the participants and appears to show an even divide, when looking at the reported answers about how participants view their own pronunciation, a comparison against the native model is present in almost all responses.

The 4 participants listed below all stated that they do not aim to sound like a native speaker, yet in response to this question about their own pronunciation, compare their own pronunciation to native speakers in some way.

P3 - *[It's] better than standard Dutch English speakers, very British and not American.*

P4 - *It's ok but I still have some trouble with vowel sounds.*

P5 - *I think it sounds like RP with a Dutch accent.*

P6 - *I think my pronunciation is good and I have had people ask me where I am from, so that is always nice to hear.*

P3 and P5 explicitly compare themselves to native varieties. Although P4 does not mention a particular English variety, by suggesting she has 'trouble with vowel sounds,' she suggests that she is not yet at the standard she could be, and her vowels are in need of improvement to become sufficiently English. P6 comments that it is complimentary when others ask where she is from as they cannot tell she is Dutch, meaning they cannot immediately tell if she is a native speaker or not.

The remaining participant who stated they do not aim to sound like a native speaker is P7. This participant is also the only one who reported that Dutch English is acceptable pronunciation if it is understandable.

P7 - *I think my own pronunciation is pretty good even though I don't speak RP or GA, but a mixture.*

The difference in this response is that P7 mixes both varieties. Her ideological view about 'good' pronunciation is reflected within her own use of the language which does not replicate one given variety, but combines aspects of both to create an 'international English.' However, even in this case P7 compared her own English pronunciation with RP and GA.

With so many varieties of English, some research has suggested adopting a standard international form as an answer to adapt to the growing number of speakers and identities to be represented worldwide (Jenkins. 2007). However, when participants were asked whether they can tell if someone is a native English speaker or not, 7 stated that they could not always tell. Out of these 7 participants, 6 responses also included a reason that second language speakers would make errors - whether it was with allophones, intonation or 'quirky pronunciation.' Only 1 participant suggested that her inability to tell was due to her lack of experience with different kinds of English pronunciation, such as Australian English. If this is the case, should aiming to sound like a native speaker be an important factor in language learning when you might not recognise if someone is native yourself?

4.2.3 Experience with Dutch second language speakers

To end the interview, participants were asked what they would think of a second language learner of Dutch with an unusual pronunciation, but who was easy to understand. The purpose of this question was to switch the perspective of the participants from second language user to first language user, and make them ask themselves if they would expect a second language learner of Dutch to sound native-like, or if their pronunciation would be considered good enough if they were understandable.

The responses were coded into the following categories: *positive*, *neutral*, *negative*. In this case, neutral responses were coded if a participant described the situation rather than their opinion of it. The results were plotted in Fig. 4.4.

Fig. 4.4: Table showing the frequency of positive, neutral and negative responses to second language learners of Dutch

Code	Positive	Neutral	Negative
Opinion of a second language speaker of Dutch with a 'strange' pronunciation	4	3	3

In comparing these results on an individual level with responses to previous questions, there appeared to be no individual consistency or category that would predict how participants reacted to a 'strange' Dutch L2 pronunciation based on their thoughts regarding English L2 pronunciation.

3 of the 4 participants who responded positively were aiming to sound like a native speaker of English themselves. In particular, P1 and P2 commented on the fact that learning a new language is hard enough without having to pronounce everything correctly. These comments are shown below.

P1 - *I'd think it's great that they try, since Dutch is a difficult language even for most Dutch people (at times).*

P2 - *It does not bother me immensely. I applaud every person trying to learn a new language or trying to pronounce everything correctly.*

Do these opinions in fact contrast with their view of their own second language endeavours? As English is much more widely spoken and international than Dutch, there are more variations in pronunciation. However, a possible reason Dutch pronunciation is seen less negatively in these cases could be that the proficiency of the imagined speaker is perceived as being low. P1 and P2 both mention that the speaker is 'trying' to learn Dutch, which will not be comparable to how they both 'use' English.

The most negative response to this question came from P10, who stated that she would ask the speaker to switch to English. However, she later compares 'bad' Dutch pronunciation to 'bad' English pronunciation, and states that the latter is more of an annoyance.

P10 - ... *but I wouldn't mind that much. It would bother me more if someone had a bad English pronunciation, which sounded really Dutch.*

The remaining 2 negative responses did not express an especially strong negative view. Both contrast what they are saying with the acceptance that not every language learner finds pronunciation as important as they do.

P6 - ... *it annoys me but only because I know better.*

P8 - *I'd probably frown a bit, but you can't expect everyone to speak without an accent.*

Overall, the responses about the pronunciation of second language Dutch users were not viewed as strongly as English second language speakers. For these participants, English is an indexical of two distinct identities - RP and GA - and even though being understandable is important for pronunciation, it is often not enough to be viewed as being 'good.'

5 Discussion

5.1 Overview

The results of this study show a clear hierarchy of preferred pronunciation among the participants. The more a pronunciation replicated RP, the more it was evaluated as 'correct'. Within section 3.2.2, a pattern emerged that no matter what category of pronunciation participants were asked to rank, the results of the mean scores were largely predictable based on the speaker being rated, rather than the category of pronunciation.

However, the results are not as simply defined as the above suggests. This discussion section will address three factors that appear to have influenced the results and changed how participants approach their evaluations of the four speakers. These factors are; the consistency in the effect of prosodic features; the familiarity or foreignness of a pronunciation; and how expectations of English are linked to perceived intelligibility.

5.2 The defining feature of good pronunciation

The results of this study are consistent with those of Derwing (2003) in finding that when participants are asked to identify specific issues or successes in pronunciation, most responses display general associations such as 'improvement of vowels' or point towards specific deviations in phonemes that are unlikely to make much difference to intelligibility. An example of such a deviation that was noted within the current study was the difference between /v/ and /w/, which was referred to in 5% of comments for the New Delhi speaker (See Fig. 3.4, section 3.2.3). This deviation is found in the speaker's pronunciation of 'very warm', which in RP would be /veri wɔ:m/, but was produced as /veri: vɔ:m/.

The deviation from RP phonemes that was most commonly referred to as feature to improve was the pronunciation of /θ/, which was produced as either /d/, /t/, or /s/. This segmental deviation was referred to for the New Delhi (13%), Groningen (7%), and Zhengzhou (8%) speakers. Other commonly mentioned deviations from RP included the difference between /l/ and /t/ (6% for the New Delhi speaker), pronunciation of /r/ (7% for the Zhengzhou speaker) and final-obstruent devoicing (7% for the Groningen speaker). Strikingly, the 'mispronunciation' of RP /θ/ and /t/ were noted by Jenkins (2007:148) as features that are not found to contribute to intelligibility of speech, and are therefore not critical for learners to replicate.

As mentioned in section 2.1, although Jenkin's (2007:148) table of pronunciation features that do not impact intelligibility are based on little empirical research, these consonants have been chosen based on the sounds that are meant to be the most difficult for second language learners to produce. It may be noted that these are in turn seen as the 'difficult' sounds to master within English, and therefore may be the most commented upon pronunciation errors because it has

become expected that a 'good' English speaker will be able to produce these consonant sounds, whereas a 'bad' speaker will not.

Final-obstruent devoicing, which is not included in Jenkin's (2007) list, is a common transfer from Dutch in which a phonological process results in voiced final-obstruents being produced as voiceless when followed by a fricative, sonorant or, in some cases, a plosive (Grijzenhout 2001:207). This devoicing results in the production homophones such as 'slip' and 'slib' which Gussenhoven & Broeders (1997:129-130) refer to as a 'frequent error' for both beginner and advanced Dutch learners of English, and is therefore likely to be a deviation from RP that the Dutch participants are very conscious of when observing another Dutch speaker's English pronunciation.

No particular phonemes stood out in the responses as a pronunciation feature to improve. All 4 speakers received comments about the pronunciation of consonants that they needed to improve, though these rarely overlapped due to the different influences that affect each accent. Views on the acceptability of correct phoneme usage seemed to be related to how 'authentic' an accent sounded overall. If participants regarded a speaker as non-native, phonemic deviations would be seen negatively. However, if the participant believed the speaker was a first language speaker, these deviations were seen as more acceptable.

In comparison, the only consonant that was mentioned as a successful pronunciation feature was the absence of the rhotic /r/, which is generally the norm of Englishes spoken in the UK (Trudgill 2010:144). In all other cases where good pronunciation referred to the production of a phoneme, all comments were general. In testing Spanish learners' evaluations of Spanish varieties, McBride (2015:28) found that the participants who were second language users of Spanish focused too exclusively on 'form errors,' rather than overall fluency of speech. This may help explain the results found above, as the majority of comments were focused on deviations from RP, or specific 'form errors' of the speakers, with no apparent consistency as to why these features were seen as correct or not when they made no difference to the intelligibility of the speakers.

However, within this study participants did consider the overall fluency of speakers. Intonation emerged as a strong divider, being remarked on as points of good pronunciation for the highly rated London and New Delhi speakers, and as points of bad pronunciation for the lower rated Groningen and Zhengzhou speakers. Although rhythm of speech was mentioned equally as a feature to improve for all speakers, it was only mentioned as a good feature for the London and New Delhi speakers. These prosodic features are the only features in the entire study which showed signs of consistency across all speakers, and which for the Zhengzhou speaker in particular, were noted by the participants as a potential indicator for unintelligibility, as the unusual intonation pattern was difficult to follow. In an accent reduction study in which participants felt their second language English accent was unintelligible and hindered communication, Behrman (2014:556) concluded that prosodic features are 'just as important as segmental ones,' and rhythmic training should be included in pronunciation teaching. Due to the difficulty of implementing prosodic training in the classroom, its importance is often seen as being secondary to other parts of the language learning process, but these results could indicate that for advanced users of English, intonation and stress are the key features that make the difference between a pronunciation being regarded as good or not.

Prosodic features, in particular intonation, appear to be the defining feature of good pronunciation. However, these results have also shown that speakers associate replication of RP features as being good, and that it was difficult for participants to state exactly what the problems with a speaker's pronunciation were without the guideline of the RP phonemic inventory to compare against.

Participants' expectations of English pronunciation appear to be strongly associated with RP. The next section will consider pronunciation expectations in more detail, and discuss how recognition of a pronunciation can influence the overall evaluation of it.

5.3 The familiarity of the pronunciations

Within section 2.3 of this study, it was suggested that a participant's familiarity with another pronunciation of English could have a positive effect on their categorisation of that pronunciation (Lindemann 2003, McKenzie 2008). However, within this study the most familiar pronunciation to the participants was the Groningen speaker, the pronunciation of which was overall categorised with an equal distribution of positive and non-positive evaluations.

On the other hand, the New Delhi speaker's pronunciation could not be identified by the majority of participants, yet this speaker received a much higher rating. Below I will discuss how the familiarity of each speaker's pronunciation is reflected in the participants' responses, and how this has influenced how each speaker's English pronunciation is rated.

5.3.1 Familiar pronunciations

The London and Groningen speakers were both geographically labelled by 40% and 66% of participants respectively, showing familiarity with these two pronunciations. Although the London speaker was never identified correctly as being from southern England, participants attempted to show recognition of where the speaker was from by labelling them to areas where English is spoken as a first language, including locations such as Ireland and Scotland, showing that they knew he was from the UK. This attempted speaker labelling occurred 6 times for the New Delhi and Zhengzhou speaker combined, showing the unfamiliarity of these pronunciations.

Although the London speaker's pronunciation was compared with RP, which led to questioning where he could be from, on the whole, participants still saw him as being a native English speaker. Therefore, although his deviations from RP were often regarded as errors, consideration of him as a 'native speaker' based on his assumed location meant that his pronunciation was overall categorised highly. In the few cases where the London speaker was not thought to be a native speaker, these same deviations lowered his overall evaluation scores. This finding may be explained by the results of Hoorn et al (2014) who found that in their pronunciation grading experiment, the one native speaker included in their sample of students was not always identified, and that in these cases they would be given a much lower grade than when they were identified as being a native

speaker, in which she was awarded full marks. The familiarity of the London speaker as British or Irish led to identification of them as a native speaker, which led to certain expectations that his pronunciation must be good.

On the other hand, when the Groningen speaker was labelled to a certain location, the speaker was always correctly identified as being Dutch, and the phonetic or prosodic deviations that were seen as being a transfer errors (note that Dutch-specific transfer errors were purposely prompted in the methodology to test this recognition. See Fig. 3.5, section 3.1.2) were the features that participants used to identify him as Dutch. Therefore, if the speaker is using 'typically' Dutch features within their English, despite the fact this is the most familiar pronunciation to participants, its obvious 'Dutchness' clearly distinguishes it from native English. The familiarity in pronunciation has in this case acted negatively on participants' evaluations of the Groningen speaker. This is supported by the results of the follow-up interviews, in which 100% of participants discussed phonetic deviations that are observed as being Dutch errors as features to improve, even without any stimuli to prompt their answers.

We may conclude that these particular participants do not find the familiarity of a Dutch pronunciation of English synonymous with being 'good' at English, as they instead relate the deviations from RP to their first language, Dutch. Rogerson-Revell (2014:139) stated that people still tend to negatively evaluate a person's language competency if they have an accent with obvious influence from their first language. The tendency to semiotically associate one pronunciation to one language to one country is still present even within English (Shuck 2006). Therefore, the mixed results on the Groningen speaker's pronunciation may be explained by participants' conflict in deciding whether a pronunciation that is completely understandable can still be classified as 'correct' English, even if it makes you think of a different language.

5.3.2 Unfamiliar pronunciations

Comprehensibility was a feature that was mentioned for both of the 'unfamiliar' pronunciations of the New Delhi and Zhengzhou speakers during the pronunciation evaluation task. This was not present in the comments about the two familiar accents. As the evaluations of the two unfamiliar pronunciations are completely different, reference to comprehensibility in both shows that the unfamiliar pronunciations are seen as more 'foreign'. Otherwise, whether the speaker was understandable or not would not have been a consideration for the participant. Even though some well recognised native English accents can be difficult to process, such as a strong Glasgow or Newcastle accent, there is a tendency to relate only second language influenced accents as being related to intelligibility (Rogerson-Revell 2014:153).

Overall, participants found the Zhengzhou speaker difficult to understand compared to the other three speakers. This resulted in the Zhengzhou speaker being labelled as a second language English learner. Gass & Varonis (1991:132) suggested this identification means any miscommunication in a second language speaker's speech is a sign of 'linguistic deficit.' Therefore, as the speaker was reported as being difficult to understand, it is viewed as their responsibility to improve because they are still in the learning process, and some participants offered suggestions to improve her

pronunciation such as imitating RP intonation and stress patterns. On the other hand, the New Delhi speaker was highly commended for being easy to understand. Although participants still commented on the speaker's segmental deviations from RP, some participants suggested that this pronunciation did not necessarily need to change to these standard variants.

However, despite some positive evaluation and acceptability of the New Delhi pronunciation, lack of familiarity had an overall negative effect on its evaluation. Although it was generally high, the mean score was never as high as the London speaker in any category due to this 'foreign' perception. The unfamiliarity of the pronunciation meant that the speaker was not identifiable to the UK or USA which are semiotically associated with English, was regarded as 'foreign' and therefore not a native speaker - meaning their results could not be as high as the London speaker. It is interesting to note that in the 3 cases in which the New Delhi speaker was identified as being from India or a former colony of the UK, the participants stated that this meant she was therefore a native speaker, and gave her the highest possible scores. In consideration of the role of English as a first language in India, Mesthrie (2010:595-600) noted that within the country, the elite of India are often associated with being 'equilingual' in English and Hindi, yet the status of native speakers of English in India worldwide is questioned due to the multilingual nature of this population. This may perhaps explain why there is much less reason for participants to identify the New Delhi speaker as being a native speaker, and thus the perception of foreignness means she is judged differently from the London speaker, who has a higher association of speaking 'English.'

5.4 Understanding pronunciation and the ideology behind the native speaker model

So far, this discussion has highlighted a tendency for participants to strongly relate to their perception of native-like pronunciation as the measuring point for their evaluation of the 4 speakers. However, within the follow-up interviews good English pronunciation was not only considered to consist of sounding native-like, but also being understood. Despite mixed responses regarding whether a foreign accent would be acceptable pronunciation or not, the majority of participants compared non-native accents to RP or GA, which stood out as the two idealised forms of spoken English. This section will now observe how being 'understandable' was referred to within the results of the pronunciation evaluation task.

Although I agree with Scheuer (2005:116) that 'foreign accent and unintelligibility are not synonymous,' figures 3.5 and 3.6 suggested that there was a three-way grouping between the 4 speakers, with a clear correlation between how native-like an accent sounded and how understandable it sounded. The first group consisted of the London speaker, which can be coded as *most native-like, highest understanding*. The second group was made up of the New Delhi and Groningen speaker, which can be coded as *questioned native-like, second highest understanding*. The final group was the Zhengzhou speaker, which can be coded as *not native-like, medium understanding*. This observed link aligns with recent literature which suggests that the less intelligible a person's pronunciation is thought to be, the more likely they are to be regarded as a second language speaker with interference from their first language (Rogerson-Revell 2014:139).

However, group 2 is the most interesting result. The New Delhi speaker has a highly preferred pronunciation, and their median native-speaker score is considerably higher than that of the Dutch speaker. As these two speakers have almost identical scores for comprehensibility (both rating highly), at first it appears that another factor must be playing a role in this positive evaluation. However, the large distribution of scores for both of these speakers, with both being given highest scores of 8 or 7 and a lowest score of 1, suggests that these two groups are in fact quite similar as both are comprehensible yet have ambiguity in their native speaker scores.

Defining the traits of native English speakers has been shown to be increasingly difficult and decreasingly important as international use of the language grows (Crystal 1997). So why does it matter that participants struggled to define whether these two speakers were native-like or not? The results throughout this study have shown that the comparison to RP phonetic and prosodic features have been the measuring stick against which all 4 speakers have been held against. Within the follow-up interviews, all participants referred to native-like pronunciation as a key part of good English pronunciation within at least one of their responses. The association of RP being the correct form by these second language English users has been shown to be exceptionally strong, despite the fact these Dutch participants have learnt English as an international communication tool outside of an English-speaking country - and not to communicate with first language speakers or assimilate to a particular culture. The second group within the three-way grouping shows that high understanding of pronunciation is not regarded as being enough to be 'native-like' if they do not represent the standard varieties. Thus, the pronunciation of these speakers is likely to be seen as a 'lower standard' in the language ideology of the participants, in which it is acceptable to reject deviations from an expected norm and justify a certain restriction of the standard language (Lippi-Green 1997:72).

5.4.1 Implications on English teaching

The conclusion of section 5.3 does not necessarily predict a future in which the two native-like pronunciations of RP and GA continue to be the idealised norms. Instead, I argue that these results show the requirement for an increase in sociolinguistic competence within English language classrooms. It is unrealistic to suppose that multiple or mixed varieties can be taught within a single English language classroom and it is certainly necessary for teachers to make an informed decision about which variety is most suitable for instruction (McBride 2015:14). The argument that RP and GA are the most suitable due to their worldwide recognition and high availability of stimuli would be unreasonable to argue against (Szpyra-Kozłowska 2015:31). However, within this study participants considered RP indexical of Britain and therefore this variety is strongly associated with being 'British' and speaking English. This association resulted in almost all participants being unable to identify the location of the London speaker - even in general terms as being from southern England - despite the fact the speaker has spent the majority of their life in the capital of the country that the participants associate their own English pronunciation with.

Speakers of RP form a relatively small part of the UK population and are associated with a certain social class, not a location (Szpyra-Kozłowska 2015:31). This is not a matter of importance to the curriculum of second language learners, and especially those who most often use English as a lingua franca. However, this study has shown the unrealistic perception of RP as the native speaker model to compare all pronunciations against, and I question the importance of pointing out segmental and prosodic deviations from this variety that is mostly limited to a small minority of speakers within the UK. In this study, comments on the deviations from RP may have arisen from the task itself, with participants seemingly considering that pointing out any deviations as incorrect forms would display their own high knowledge of English. Instead, this information has evidenced that knowledge of linguistic diversity is missing from English second language education.

At the present time, the most important benefit of speaking English for most learners is to be able to communicate with as many people as possible - both first and second language users (van den Doel 2006:4). Pronunciation is an undeniable part of second language instruction, from stimuli in the classroom to students' own production, and these results have shown that sounding like a particular inner circle, native speaker is still considered as being important for this set of future instructors. However, 'recognising that language variation is a natural part of communication - and thus, necessary for communicative competence' would not only increase second language users' familiarity with and ability to speak with a variety of international English users, but would also increase the acceptance of distinct English language pronunciations by this group (Geeslin & Long 2014:258). If RP and GA continue to be the primary pronunciation models for future English students despite the growing use of the language as an international lingua franca, there is an overwhelming need to increase the use of other pronunciation models as stimuli to address the practical realities of using English as a second language today.

6 Conclusion

6.1 Overall findings

The purpose of this study was to discover how second language users of English, in particular Dutch students who major in English Language & Culture and expect to become future English teachers themselves, evaluate the ‘correctness’ of varieties of global Englishes that reflect the current diversity in the use of the language worldwide, and discover what the important factors are in determining what makes ‘correct’ English pronunciation. The results hoped to give an insight into how impactful the native-speaker model is in second language teaching, as pronunciation is an undeniable part of how a person is perceived, and one that students learn through the stimuli they interact with. Below I will address the research questions set out in section 1.2.

R1: *Do Dutch students majoring in an English Language & Culture evaluate first language speakers’ pronunciation of English more positively than second language speakers’ pronunciation of English?*

I hypothesised that participants would evaluate the speakers with the least amount of deviation from the standard taught norms of RP or GA as being the most correct, predicting that the two first language speakers would be seen more positively than the two second language speakers. This was confirmed as the results of the pronunciation evaluation task showed a clear hierarchy of preferred pronunciation, with the L1 London and New Delhi speakers being overall categorised positively by over 100% and 85% of participants, respectively, whereas the second language speakers were categorised positively by 50% of participants or less (Fig. 3.6, section 3.2.1). In both the pronunciation evaluation task and follow-up interviews, L2 interference was in almost all cases categorised negatively and related to specific consonant and vowel deviations from RP.

However, when looking at the factors that prompted these evaluations, a more complex picture arose than the results first suggested.

R2: *What are the most important factors in determining the ‘correctness’ of English pronunciation?*

Prosodic features such as intonation and rhythm were the only features that divided the preferred pronunciations of the London and New Delhi speakers to the dispreferred pronunciations of the Groningen and Zhengzhou speakers, the former being said to produce ‘good’ intonation and the latter too much interference from their first languages. It was found that even if the speaker produced phonemes that deviated from an RP pronunciation, such as in cases where the London speaker was not identified as a native speaker or the New Delhi speaker was labelled as ‘foreign,’ if the speaker used intonation that showed no interference from another language, these phonemic deviations were more easily accepted. This result supports Mesthrie’s (2010) conclusion that prosodic features are just as important as segmental features in pronunciation, and I would even suggest that for advanced learners, intonation is the defining point of successful pronunciation. Future studies are required to empirically test this claim, and could use matched-guise tests in which speakers have artificially modified intonation to assess how much of a difference this makes in the

intelligibility and overall perceived 'correctness' of a variety.

Two further sociolinguistic reasons were found to influence how the participants evaluated the speakers' pronunciations. Firstly, if the speaker was labelled as having a pronunciation that was deemed as 'foreign' from RP or GA, the pronunciation would be defined as substandard regardless of how intelligible the speaker was. This was especially evident in the case of the Groningen speaker, who was the most identified pronunciation and well understood, yet whose noticeably Dutch features were evaluated negatively. This finding was in contrast to previous literature which suggested that the more familiar a person is with an accent, the more likely they are to give it a positive evaluation (Lindemann 2003; Gass & Varonis 1984). A potential reason for this contrast may be that in the present study, participants were asked to define what was the 'correct' pronunciation, whereas past literature has focused on whether particular identities are systematically associated with certain accents. In this case, what was correct was associated with what was RP or GA, no matter how well understood the speaker was.

Secondly, it was found that in the cases where a speaker was easily understandable yet did not represent an accent that was either RP or GA, participants gave extremely contrasting scores regarding to what degree they thought a speaker was a native speaker. This supports the finding that English pronunciation is highly indexical of RP or GA, and calls to attention the need for increased variation in teaching stimuli so that students are aware of not only the type of English they are learning, but of a multitude of varieties so that they are equipped to interact with as many speakers as possible.

6.2 Implications for future research

This study has focused on users of English in the Netherlands, the country with the highest current English language proficiency rating in the world (EFI 2018), and more specifically on English Language & Culture students. To gain a full picture of how second language users define successful pronunciation, future research could expand to include users from other countries with more diverse backgrounds. As the majority of students in the Netherlands learn RP pronunciation, this study has focused on 4 speakers of global Englishes that either learnt RP themselves or whose education system has a current or historical connection to the UK. It would also be interesting to expand this research and see if the same clear hierarchy of preferred pronunciation appears for pronunciations that are instead influenced by GA, or if there is a difference in how GA-influenced Englishes are compared against RP-influenced Englishes.

Communicative effectiveness in any exchange relies on both the speaker and listener. In second language exchange, it has been noted that the more native-like the conversational participant appears to be, the less likely the burden of responsibility will fall on them if miscommunication arises, whereas in first language talk any miscommunication will be the responsibility of both parties (Lippi-Green 1997:69). The evaluations of second language users of English as listeners are important to consider because if the acceptability of global Englishes is to increase, the ways in which the listener responds to perceived 'correct' or 'incorrect' English will help improve cross-

communication barriers.

Recent literature regarding English language teaching norms has focused heavily on the potential implications of English as the current global lingua franca, and this has inspired many studies on the topic of future English norms, with some studies debating the need for a specifically built curriculum for the purposes of ELF (Jenkins 2007; Seidlhofer 2011; Canagarajah 2007). In response, I suggest that it is unrealistic and of little benefit for the future of ELF to rebuild an established English language curriculum to a description of alternate norms that cannot possibly be adapted to every linguistic background. Instead, the results of the present study have shown that there is a need to increase the focus on parts of the language curriculum that are not currently adhered to - those being; the importance of prosodic features; and student familiarity with the diverse population of English language users that they will soon be joining, to raise awareness of how pronunciation can be distinct and represent cultural identities without having to be deemed as incorrect or less intelligible. These results provide empirical support for Geeslin & Long's (2014:258) theory that basic knowledge of language variation in a person's second language is increasingly important to ensure that user gains communicative effectiveness with as many other speakers as possible, which is, after all, the primary aim of most people when learning a second language.

No matter how much pronunciation training is or is not used in the classroom, is it undeniable that through the stimuli and language that students are exposed to, pronunciation matters. Taking into account the current status of English worldwide, the association of English being indexical of one or two varieties, that are in practice used by a minority of speakers, should be restructured to give learners exposure to a wider variety of English speakers around the globe, for the uses and teaching of English as an international language to flourish.

References

Articles / Books:

- Abeywickrama, P. (2013) Why not non-native varieties of English as listening comprehension test input? In: *RELC*. Vol 44 (1). pp:59-74
- Behrman, A. (2014) Segmental and prosodic approaches to accent management. In: *American Journal of Speech-Language Pathology*. Vol 23. pp.546-561
- Canagarajah, S. (2007) Lingua Franca English, Multilingual Communities, and Language Acquisition. In: *The Modern Language Journal*. Vol 91. pp.923-939
- Cogo, A. (2012) English as a Lingua Franca: concepts, use, and implications. In: *ELT Journal*. Vol 66(1). pp.97-105
- Collins, B., Hollander, S., & Rodd, J. (1987) *Accepted English Pronunciation: A practical guide for speakers of Dutch*. Apeldoorn: Van Walraven
- Collins, B., den Hollander, S., Mees, S., & Rodd, J. (2011). *Sounding better: A practical guide to English pronunciation for speakers of Dutch*. Holten: Walvaboek
- Crystal, D. (1997) *English as a global language*. Cambridge: Cambridge University Press
- Cutler, C. (2014). Accentedness, "Passing," and Crossing. In J. Levis & A. Moyer (Eds.), *Social dynamics in second language accent*. pp. 145–167. Boston: DeGruyter.
- De France, T., & Smakman, D.(2013) The effects of teaching Dutch learners British English consonants. An experimental study. In: Dick Smakman & Laurent Willemsen (Eds.) *Proceedings of the Van Schools tot Scriptie I Colloquium*, Leiden, 8 & 9 June 2012
- Derwing, T. (2003) What do ESL students say about their accents? In: *The Canadian Modern Language Review*. Vol 59(4). pp.547-566
- Derwing, T., & Munro, M. (2009) Putting accent in its place: Rethinking obstacles to communication. In: *Language Teaching*. Vol 42(4). pp.476-490
- Dragojevic, M., Giles, H. (2016) I don't like you because you're hard to understand: the role of processing fluency in the language attitudes process. In: *Human communication research*. Vol 42. pp.396-420
- Dragojevic, M., Mastro, D., Giles, H., & Sink, A. (2016) Silencing nonstandard speakers: a content analysis of accent portrayals on American primetime television. In: *Language in society*. Vol 45. pp.59-85
- Dragojevic, M., Giles, H., Beck, A., & Tatum, N. (2017) The fluency principle: Why foreign accent strength negatively biases language attitudes. In: *Communication Monographs*. Vol 84(3). pp.385-405

- Fraser, C., & Kelly, B. (2012) Listening between the lines: Social assumptions around foreign accents. In: *Australian Review of Applied Linguistics*. Vol 35(1). pp.74-93
- Friedrich, P. (2000) English in Brazil: functions and attitudes. In: *World Englishes*. Vol 19. pp.215-224
- Galloway, N., & Rose, H. (2015) *Introducing Global Englishes*. London: Routledge
- Gass, S., & Varonis, E. (1991) Miscommunication in nonnative speaker discourse. In: Coupland, N., Giles, H., & Wiemann, J. (ed) *"Miscommunication" and problematic talk*. London: Sage Publications Ltd
- Gass, S., & Varonis, E. (1984) The effect of familiarity on the comprehensibility of nonnative speech. In: *Language learning*. Vol 34. pp.65-89
- Geeslin, K., & Long, A. (2014) *Sociolinguistics and second language acquisition*. New York: Routledge
- Grijzenhout, J. (2001) Voicing and devoicing in English, German, and Dutch; evidence for domain-specific identity constraints. In: Dziubalska-Kolaczyk, K. (ed.) *Trends in Linguistics. Studies and Monographs 134: "Constraints and Preferences"*. pp.201-228. Berlin: Mouton de Gruyter
- Gussenhoven & Broeders (1997) *English pronunciation for student teachers*. Groningen: Wolters-Noordhoff
- Haslam, M., & Zetterholm, E. (2016) Vowel length and perception of English as a Lingua Franca. In: *Journal of the Acoustical Society of America*. Vol 139(4)
- Hendriks, B., van Meurs, F., & de Groot, E. (2017). The effects of degrees of Dutch accentedness in ELF and in French, German and Spanish. In: *International Journal of Applied Linguistics*. Vol 27(1). pp.44-66
- Hilton, J. (2015) *Lost Horizon*. London: Vintage
- Hoorn, M., Smakman, D., & Foster, T. (2014) Are we at sixes and sevens? Pronunciation assessment by lecturers of English in the Netherlands In: van den Doel, R. Rupp, L. (ed) *Pronunciation matters: Accents of English in the Netherlands and elsewhere*. pp.95-108. Amsterdam: VU University Press
- Hu, G., & Lindemann, S. (2009) Stereotypes of Cantonese English, apparent native/non-native status, and their effect on non-native English speakers' perception. In: *Journal of Multilingual and Multicultural Development*. Vol 30(3). pp.253-269
- Jenkins, J. (2009) *World Englishes*. Oxon: Routledge
- Kachru, B. (2009) Asian Englishes in the Asian Age: Contexts and Challenges. In: Murata, K., & Jenkins, J. (ed) *Global Englishes in Asian Contexts: Current and Future Debates*. Great Britain: Palgrave Macmillan
- Kirkpatrick, A. (2007) *World Englishes: Implications for International Communication and English Language Teaching*. Cambridge: Cambridge University Press.
- Lindemann, S. (2003) Koreans, Chinese or Indians? Attitudes and ideologies about non-native English

speakers in the United States. In: *Journal of Sociolinguistics*. Vol 7(3). pp.348-364

Lippi-Green, R. (1997) *English with an accent*. London: Routledge

Major, R., Fitzmaurice, S., Bunta, F., & Balasubramanian, C. (2005) Testing the effects of regional, ethnic, and international dialects of English on listening comprehension. In: *Language Learning*. Vol 55. pp.37-69

Margić, B. (2017) Communication courtesy or condescension? Linguistic accommodation of native to non-native speakers of English. In: *Journal of English as a Lingua Franca*. Vol 6(1). pp.29-55

Mesthrie, R. (2010) New Englishes and the native speaker debate. In: *Language Sciences*. Vol 32. pp.594-601

McBride, K. (2015) Which features of Spanish learners' pronunciation most impact listener evaluations? In: *Hispania*. Vol 98(1). pp.14-30

McKenzie, R. (2008) Social factors and non-native attitudes towards varieties of spoken English: a Japanese case study. In: *International Journal of Applied Linguistics*. Vol 18(1). pp.63-88

Pitzl, M. (2015) Understanding and misunderstanding in the Common European Framework of Reference: what we can learn from research on BELF and intercultural communication. In: *JELF*. Vol 4(1). pp.91-124

Rogerson-Revell, P. (2014) Pronunciation Matters: Using English for international business communication In: van den Doel, R. Rupp, L (ed) *Pronunciation matters: Accents of English in the Netherlands and elsewhere*. pp.137-157. Amsterdam: VU University Press

Scheuer, S. (2005) Why native speakers are (still) relevant. In: Dziubalska-Kołodziej, K., & Przedlacka, J. (ed) *English pronunciation models: a changing scene*. Berne: Peter Lang.

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

Sewell 2013, A. (2013) English as a Lingua Franca: Ontology and Ideology. In: *ELT Journal*. Vol 67(1). pp.3-10

Shuck, G. (2006) Racializing the nonnative English speaker. In: *Journal of Language, Identity & Education*. Vol 5. pp.259-276

Smith, L., (2009) Dimensions of understanding in cross-cultural communication. In: Murata, K., & Jenkins, J. (ed) *Global Englishes in Asian Contexts: Current and future debates*. Great Britain: Palgrave Macmillan

Sowden, C. (2012) ELF on a mushroom: the overnight growth in English as a Lingua Franca. In: *ELT Journal*. Vol 66(1). pp.89-96

Starks, D., & Paltridge, B. (1996) A note on using sociolinguistic methods to study non-native attitudes towards English. In: *World Englishes*. Vol 15. pp.217-224

Sung, C. (2014) Global, local or glocal? Identities of L2 learners in English as a Lingua Franca communication. In: *Language, culture and curriculum*. Vol 27(1). pp.43-57

Szpyra-Kozłowska, J. (2015) *Pronunciation in EFL Instruction: A Research-Based Approach*. Great Britain: Lavenham Press Ltd

Trudgill, P. (2010) "The new non-rhotic style." In: *Investigations in Sociohistorical Linguistics: Stories of Colonisation and Contact*. (ed). pp.143-158. Cambridge: Cambridge University Press

Trudgill, P. (2005) Native speaker segmental phonological models. In: Dziubalska-Kotaczyk & Przedlacka (2005). pp.77-98

Coetzee-Van Rooy, S. (2009) Intelligibility and perceptions of English proficiency. In: *World Englishes*. Vol 28. pp.15-34

van den Doel, W. (2006) *How friendly are the natives? An evaluation of native-speaker judgements of foreign-accented British and American English*. Utrecht: LOT

Digital resources:

Boersma, P., & Weenink, D. (2018) Praat: doing phonetics by computer. Version 6.0.37, retrieved 1 February 2018 from <<http://www.praat.org/>>

EFI English Proficiency Index, accessed 30 April 2018. <<https://www.ef.nl/epi/>>

Qualtrics, hosting the online survey and accessed throughout February and March 2018. <<https://www.qualtrics.com/>>

Appendices

Appendix 1: Link to the Pronunciation Evaluation Task and recordings

Evaluation Task <https://leidenuniv.eu.qualtrics.com/jfe/form/SV_5BhSuyfb6ursKsl>

Recordings <<https://soundcloud.com/amy-cat-609299371>>

Screenshot showing the interface of the Evaluation Task:

Please listen to the following short audio clip of speaker 'green.' Once you have listened to the whole clip, please answer the questions below. The speaker has been provided with random excerpts from a book. Focus on listening to their pronunciation of English.



Appendix 2: Full details of the outcomes of the post-hoc statistical tests from in Tables 1, 2 (section 3.2.2) and table 3 (section 3.2.6) below

Table 1: Comparison of the results of the Wilcoxon signed rank post-hoc tests comparing each speaker for each Likert scale in section 3.2.2. The blue boxes mark the instances in which a speaker was found to be significantly different from another speaker.

Comparisons	Overall Pronunciation		Intonation		Use of correct vowels		Use of correct consonants		Accent consistency	
	Z	p	Z	p	Z	p	Z	p	Z	p
London - New Delhi	3.29	.001	2.13	.033	3.90	.000	3.61	.000	1.49	.136
London - Groningen	4.49	.000	3.55	.000	4.53	.000	4.49	.000	2.30	.021
London - Zhengzhou	4.91	.000	4.24	.000	4.65	.000	4.66	.000	3.50	.000
New Delhi - Groningen	2.22	.026	2.03	.043	2.40	.018	1.85	.097	9.38	.402
New Delhi - Zhengzhou	3.94	.000	4.56	.000	4.00	.000	3.17	.000	2.72	.006
Groningen - Zhengzhou	3.21	.001	4.07	.000	2.07	.038	2.00	.045	1.79	.074

Table 2: The complete results of the Wilcoxon signed rank post-hoc tests comparing scores of ranks for individual speakers in section 3.2.2. The London speaker had no significant difference between scores, so post-hoc tests were only completed for the New Delhi, Groningen and Zhengzhou speakers. The blue boxes mark the instances in which there was a significant difference between scores for Likert scales for an individual speaker.

Comparisons between different Likert scales	New Delhi		Groningen		Zhengzhou	
	Z	p	Z	p	Z	p
Pronunciation - Intonation	1.65	.099	2.96	.003	3.38	.001
Pronunciation - Correct Vowels	1.34	.182	3.34	.001	.503	.615
Pronunciation - Correct Consonants	3.08	.020	2.92	.003	.874	.382
Pronunciation - Accent Consistency	-2.04	.041	-1.63	.104	-3.29	.001
Intonation - Correct Vowels	-5.37	.591	.229	.819	-3.46	.001
Intonation - Correct Consonants	1.18	.239	-.412	.680	-2.56	.005
Intonation - Accent Consistency	-3.33	.001	-3.58	.000	-4.56	.000
Correct Vowels - Correct Consonants	2.49	.012	-.785	.433	.627	.531
Correct Vowels - Accent Consistency	-3.09	.002	-3.85	.000	3.37	.001
Correct Consonants - Accent Consistency	-3.30	.001	-3.82	.000	-3.29	.001

Table 3: Showing full significance levels from the Wilcoxon signed rank post-hoc tests for the native speaker question in section 2.3.6.

				Test Statistics ^{a,c}					
				NN NS Rank - LDN NS Rank	GG NS Rank - LDN NS Rank	ZZ NS Rank - LDN NS Rank	GG NS Rank - NN NS Rank	ZZ NS Rank - NN NS Rank	ZZ NS Rank - GG NS Rank
Z				-4,282 ^b	-4,580 ^b	-4,848 ^b	-2,817 ^b	-4,463 ^b	-2,815 ^b
Asymp. Sig. (2-tailed)				,000	,000	,000	,005	,000	,005
Monte Carlo Sig. (2-tailed)	Sig.			,000	,000	,000	,005	,000	,004
		99% Confidence Interval	Lower Bound	,000	,000	,000	,003	,000	,003
			Upper Bound	,000	,000	,000	,006	,000	,006
Monte Carlo Sig. (1-tailed)	Sig.			,000	,000	,000	,002	,000	,003
		99% Confidence Interval	Lower Bound	,000	,000	,000	,001	,000	,001
			Upper Bound	,000	,000	,000	,003	,000	,004

a. Wilcoxon Signed Ranks Test

b. Based on negative ranks.

c. Based on 10000 sampled tables with starting seed 957002199.

Appendix 3: List of questions asked in the follow-up interviews. (Section 4.2)

1. How would you describe good English pronunciation?
2. What do you think is the most important feature of good English pronunciation?
3. Do you think there is more than one type of good English pronunciation? (with varieties such as RP, GA, etc.) If so, what types of English would you consider to be good?
4. Can you give one example of a feature that will make a speaker have bad English pronunciation?
5. Can you always tell if someone is a native speaker of English or not?
If so, how can you tell? If not, why do you think you can't always tell?
6. Is it your aim to sound like a native speaker of English?
7. What do you think about your own pronunciation of English and why?
8. What would you think if you met a second language learner of Dutch who you can understand, but they have a strange pronunciation?