Accents in Context

Non-native and native listeners' perceptions
and understanding of Dutch-accented English

Published by
LOT
Kloveniersburgwal 48
1012 CX Amsterdam
The Netherlands

phone: +31 20 525 2461

e-mail: lot@uva.nl
http://www.lotschool.nl

Cover illustration, editorial design: Nicolet Pennekamp

ISBN: 978-94-6093-337-0
NUR: 616

Accents in Context

Non-native and native listeners' perceptions and understanding of Dutch-accented English

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op vrijdag 31 januari 2020
om 10:30 uur precies

door

Warda Nejjari
geboren op 18 maart 1980
te Nijmegen, Nederland

Promotoren	prof. dr. M. Gerritsen prof. dr. R.W.N.M. van Hout
Copromotor	dr. B.C. Planken
Manuscriptcommissie	prof. dr. M. van Oostendorp prof. dr. H. Van de Velde (UU) prof. dr. E. Zenner (KU Leuven, België) dr. R. van den Doel (UU) dr. B.C. Hendriks

CONTENTS

1 GENERAL INTRODUCTION

Since the beginning of the 20th Century, as part of the inheritance of the British Empire and the growing worldwide influence of the United States, English has slowly gained the global status of the most important language used in international contexts in domains of media, politics, tourism, education, academia, and business (Rosen, 2010; Crystal, 2009; Nelson, 2011; Seidlhofer, 2005; Kachru, 1983, 1992). English has been projected to reach approximately two billion speakers by 2020 (British Council, 2013). What makes English unique is that the majority of speakers of English in the world do not speak it as their first or native (L1) language, but as a second (L2) or non-native language (Crystal, 2003, 2009; British Council, 2013). L2 English speakers use English mostly to communicate with other L2 English speakers (e.g. Crystal, 2003; Nelson, 2011; Seidlhofer, 2005; Kachru, 1983, 1992). The worldwide increase in use of English by various L2 English speakers is a new linguistic reality which might affect how L1 and L2 English speakers perceive and understand L2 Englishes. This thesis aims to understand whether an L2 English accent, in this case Dutch-accented English, compared to L1 English accents, impacts listeners' speech understanding and their perceptions of speakers.

Traditionally, L1 English speakers have been defined as speakers from nations in which English is the majority's L1 language, for example, the U.K., U.S. Ireland, Canada, Australia, New Zealand, and South Africa (e.g. Crystal, 2003, 2009; Kachru, 1983, 1992). L2 English speakers have traditionally been defined as either speakers of English from former British Empire colonies in Africa and Asia, such as Singapore, Nigeria, and India, where English has an official status, or speakers of English who learn English as a foreign language in countries that were not part of the British empire, such as Germany, Belgium, The Netherlands, Sweden, Brazil, Russia, Morocco, and France (e.g. Varonis & Gas, 1982; Cargile & Giles, 1997; Kalin, Rayko & Love, 1980; Lindmann, 2002; Major, Fitzmaurice, Bunta & Balasubramanian, 2005; Munro, Derwing & Morton, 2006). Inhabitants from the latter group may also be separately categorized as speakers of English as a foreign language (FL) (e.g. Abu-Rabia, 2004; Albrechtsen, Hendricksen, Mees & Poulsen, 1998), exactly because they were not part of the British empire and therefore historically have not been exposed intensely to English in their societies. As a result they have not developed their own (official) English language varieties. However, since most distinguish only between L1 and L2 English, this thesis will do the same.

English language teaching organizations, such as the British Council, a global supplier of English testing and education, are acutely aware of the fact that English 'no longer belongs' to L1 English speakers (Crystal, 2003: p.2). They view the global spread and use of English as an opportunity to ensure that English can flourish

internationally with great variety. However, the British Council emphasizes that in order to ensure effective communication English should be taught using an L1 English variety as the main teaching foundation, or as is stated in the British Council's 2013 report: 'There is a need for high-quality teaching of English even more than ever.... to ensure that the diverse dialects that are being allowed to flourish are clear about the source code from which they are diverging and share enough common elements to 'interoperate' as a global operating system.' (British Council, 2013: p.7).

Indeed, English language education for L2 English speakers more generally has tended to be oriented towards achieving English language fluency in an L1 English variety. Two L1 English varieties that are commonly taught on a global scare are: British and American English (ETS, 2019; Cambridge, 2019; British Council, 2013). This approach is assumed to increase the chances of mutual understanding between different speaker groups of L2 English in the areas of grammar, vocabulary, and also in terms of pronunciation (see Derwing & Munro, 2009 for a summary of views on accent training for L2 English learners). We know, however, that achieving a native-like pronunciation or accent is one of the hardest language skills to master as an L2 learner, especially after the critical period, which ends around the time puberty begins (Lenneberg, 1967; Bongarets, van Summeren, Planken & Schils, 1997; Friedman & Rusou, 2015). Therefore, it is likely that L2 English speakers will always sound non-native, despite their best efforts in achieving a native norm in education (Bongaerts et al., 1997; Vermeulen & Kellerman, 1998; van den Doel, 2006). This could have consequences for L2 English speakers, because an L2 English accent can significantly impact the perceptions people have of speakers as well as their understanding of speech (e.g. Cargile & Giles, 1997; Munro & Derwing, 1995 a, b; Kalin Rayko & Love, 1980; Lindemann, 2002; Phiko, 1997). In other words, L2 English accents can impact *speaker evaluations* and *speech understandability*.

However, if English is no longer 'owned' by L1 English speakers and is spoken and used mostly by L2 English speakers on a global scale, it might be the case that L2 English speakers have adapted to the state of English as a global lingua franca to such an extent that they have become capable of understanding a variety of L2 Englishes and have become accepting of different L2 English varieties. However, despite the L2 English speakers' dominance in numbers, most English education is still aimed at teaching L1 English varieties and many L2 English speakers will still communicate in English with L1 English speakers. In order to form a more complete picture of the impact of L2 English accents on understandability and evaluation of its speakers, it is important to gain more insight into the effect of L2 English accents compared to L1 English accents on both L1 and L2 speakers of English.

Accents – such as Dutch English – are not produced in isolation but in specific communication contexts, for example, by a speaker in a lecture or in a job interview. A specific communication context can carry with it expectations about how people should generally behave, and how they should speak (Burgoon, 1979; Burgoon & Burgoon, 2001). This suggests that responses to English accents might be relative or context-dependent. Since responses to accents may be a reflection of the extent to which a speaker with a specific accent was able to meet listeners' expectations within a specific communication context and since few accentedness studies have considered multiple contexts, *communication context* is included as a variable in the present thesis (See the third and fourth study reported in Chapters 4 and 5).

This thesis is aimed at investigating whether an L2 English accent, compared to an L1 English accent, impacts *speaker evaluations* and *speech understandability*. Dutch-accented English was selected as an example of an L2 English accent. Dutch-accented English has been studied in terms of its grammatical, phonetical and phonological features as well as how specific speech features might arouse negative responses in listeners (e.g. van den Doel, 2006; Gussenhoven & Broeders, 1997; Collins & Mees, 2003; Edwards, 2016). However, to the best of our knowledge, it has never been studied what the effects of Dutch English are compared to commonly taught L1 English accents on *speech understandability* and *speaker evaluations* of different L1 and L2 English listeners in a variety of communication contexts. Therefore, in the present thesis, the responses of different listener groups (British, Dutch, German, Singaporean and Spanish listeners) to Dutch-accented English are compared to two commonly taught L1 English accent varieties: standard British and standard American English accents (ETS, 2019; Cambridge, 2019; British Council, 2013).

Before the overview and aims of the four studies conducted in this thesis are provided, the motivation for selecting Dutch-accented English as an L2 English variety is provided in section 1.1, by considering the status of English in the Netherlands and the features of Dutch-accented English.

1.1 ENGLISH IN THE NETHERLANDS AND DUTCH ENGLISH

In the Netherlands, English is the most important foreign language that is learned and used in primary, secondary, and tertiary education. Mostly, Dutch learners of English are taught L1 English varieties and the variety of English that is mostly taught is British English, with American English having gained popularity as well (e.g. van der Haagen, 1998). A growing number of secondary school pupils enroll in

bilingual English and Dutch schools, and a small number of primary school schools are part of a pilot for bilingual primary programmes in which students at the age of four receive at least a third of their education in English (Nuffic, 2019). Pupils and students are also encouraged to obtain internationally acknowledged certificates in English, for instance, the British English certificate in English, the 'Cambridge Certificate' (Rijksoverheid, 2019; Cambridge Assessment English, 2019), because such a certificate may be needed or is useful for enrollment in specific degree programmes, and may increase chances of gaining employment. The strong focus on developing good English skills illustrates how in the Netherlands having good English skills is seen as a fundamental part of a good education and increases people's chances to succeed professionally. In addition, English is also important in other domains in Dutch society as it is commonly used in the Dutch media, advertising, and business (e.g. Gerritsen, Korzilius, van Meurs & Gijsbers, 2000; Gerritsen, Nickerson, Van den Brandt, Crijns, Dominguez & van Meurs, 2007; De Groot, 2008; Gerritsen & Nickerson, 2004).

The omnipresence of English in almost all educational contexts, media, business, and academia is not always evaluated positively. It has led to discussions on the development and status of English in the Netherlands as well as debates on the English language skills of the Dutch and Dutch public figures (e.g. Gerritsen, van Meurs, Planken et al., 2016; Edwards, 2016; Onze Taal, 2014). For example, English is viewed by some as a threat to the status of other foreign languages in Dutch secondary education, such as French or German. Some even regard the status of English in the Netherlands as a threat to Dutch, for example due to the growing number of university degree programmes being offered in English. The assumption is that if an increasing number of Dutch students choose to take degree programmes in English rather than Dutch, they might only develop academic skills in English and not in Dutch (e.g. BON, 2017; Bouma, 2016, 2018; Edwards, 2016; Huygen, 2017; Lizzini, Martijn, Munk, & De Regt, 2017; van Gaal, 2018; van Heest, 2018). Furthermore, Dutch media have frequently featured debates on public figures' bad English skills and pronunciation. In 2014, for instance, the Dutch national language association 'Onze Taal' even conducted an online poll amongst its members to enquire about the state of Dutch prime minister Rutte's English pronunciation and his typical Dutch English accent, because it had been heavily scrutinized on Dutch social media (Onze Taal, 2014). The poll indicated that 55% of the approximately 800 'Onze Taal' members that were surveyed believed he should try to improve his pronunciation, that is, make his pronunciation sound less Dutch, and 45% believed it was alright for him to sound Dutch when speaking English.

While the Onze Taal members appear to have a clear idea of what a typical Dutch English accent sounds like, judging by their verdict on prime minister Rutte's accent, it is less clear from the literature what a Dutch-accented English actually is. There is no single definition of what a typical Dutch English accent is, but it is generally described as possessing phonetic features that L1 speakers of Dutch and others familiar with Dutch and Dutch English will recognize as such. One well-known example is the fact that Dutch does not have dental consonants [ð] as in *the, brother* and [θ] as in *think, Perth*. These dental consonants are often pronounced as stop consonants, [d] and [t] respectively, by Dutch speakers of English (see also Gussenhoven & Broeders, 1997; Collins & Mees, 2003). Even though there is no single definition of what a typical Dutch English accent is, and the members of a national Dutch language society might not even be a good representation of the average Dutch person, the fact that the Dutch prime minister's English pronunciation led to a national online debate, and to the abovementioned poll, illustrates the controversiality of Dutch-accented English and that the Dutch appear to be divided on the issue of the importance of sounding native-like in English, at least when it comes to those who represent Dutch interests on an international scale.

The criticisms offered by the Dutch on their own accent variety in English perhaps also reflect a language norm that views successful English language acquisition on the part of L2 learners as being able to match the language skills of an L1 English speaker. This language norm resembles the traditional second language acquisition (SLA) perspective, which is based on the idea that L1 language varieties are used or owned by L1 speakers and are learned by L2 speakers. L2 speakers are expected to achieve proficiency in a defined 'target language', an L1 language variety, and to minimize the influence of their L1 on that target language. Acquisition is considered to be most successful when the L2 speaker is able to use the target language like an L1 speaker of that language (e.g. Corder ,1967; Selinker, 1972; Richards, 1974; Ellis, 1994; Gass & Selinker, 1994; Lightbown & Spada, 1999; Penning de Vries, Cucchiarini, Strik & van Hout, 2019).

Even though the traditional SLA perspective is common in SLA research, it has been criticized for not considering the social nature of language use as well as changing sociolinguistic circumstances, which are factors that might affect language norms and perceptions of language varieties and speakers. As mentioned earlier, in the context of English as a global language, most speakers of English in the world are L2 speakers. They bring to the table their own unique cultural and linguistic identities, and a majority of them use English to interact with other L2 English speakers, with possibly other cultural and linguistic backgrounds. Therefore, they

might view English as a practical communication tool to achieve their objectives with, and as a result, L2 English accents might not necessarily be viewed as a hindrance to effective communications nor be important signifiers of someone's personal characteristics. This might mean that L2 English speakers form a new global speech community, as noted by Canagarjah (2007), who refers to this community as "the Lingua Franca English (LFE) speech community". This assumption forms the backdrop for the research presented in the present thesis. By tapping into L1 and L2 English listeners' perceptions of L1 and L2 English varieties, it aims to throw light on the idea of a potential LFE speech community and the linguistic reality of English as a global language and lingua franca.

1.2 OVERVIEW OF THIS THESIS: AIMS, VARIABLES, AND METHODS

The aim of this thesis is to understand whether an L2 English accent compared to an L1 English accent impacts the understanding of speech and the perceptions of speakers. Three studies were conducted in which an L2 English accent, namely Dutch-accented English, was compared to two L1 English accents, namely a standard British accent (studies 1,3,4) and a standard American English accent (studies 3-4). The second study was conducted to develop a reliable research method that allows for the creation of representative matched guises for the final two studies. The responses to the selected accents were assessed in terms of the *speech understandability* and *speaker evaluations* of the accents. In study 1, L1 English listeners were used, and in studies 2, 3, and 4 both L1 and L2 English speakers acted as listeners. In studies 3 and 4, the listeners responded to the tested accents in different communication contexts. The findings overall allowed us to reflect on traditional second language acquisition research perspectives, the potential existence of an international Lingua Franca English speech community, and what this could mean for English language education.

Sections 1.3-1.6 summarize, in chronological order, the aims of, and research methods of the four studies. Table 1 shows the (in)dependent variables, the listener groups, and the research techniques that were applied to create the stimuli, in each of the four studies.

Study 1 was conducted to understand to what extent sounding native-like as a Dutch speaker of English affected the responses of L1 English speakers. Highly educated L1 British English speakers, in a sales communication setting, reacted to slight and

Table 1. Independent variables, dependent variables, methods studies 1-4.

Study	Independent variables			Dependent variables		Method ***
	Accent*	**Context**	**Listeners****			
1	BrE	Telephone Sales pitch	BrE-F (n=72)	*Speech Understandability:*	Speaker Evaluations:	VG
	Slight DE		BrE-NF (n=72)	Intelligibility	Status	
				Comprehensibility	Affect	
	Moderate DE			Interpretability		
2	BrE	--	L1-BrE (n=40)	*Representativeness:*		MG
				Nativeness		
	AmE		L1-AmE (n=40)	Standardness		
	DE		L1-NL (n=40)			
3	BrE	Lecture	L1-NL (n=392)	*Speech Understandability:*	Speaker Evaluations:	MG
	AmE	Audio Tour		Intelligibility	Status	
				Comprehensibility	Affect	
	DE	Job Pitch		Interpretability	Dynamism	
4	BrE	Lecture	L1-Ger (n=617)	*Speech Understandability:*	Speaker Evaluations:	MG
	AmE	Audio Tour	L1-Spa (n=540)	Intelligibility	Status	
				Comprehensibility	Affect	
	DE	Job Pitch	L1-Sing (n=542)	Interpretability	Dynamism	

* Accents:
BrE = standard British English; Slight DE = slight Dutch-accented English;
Moderate DE = moderate Dutch-accented English; AmE = standard American English;
DE = typically Dutch-accented English of highly educated Dutch speakers.

** Listeners:
BrE-F = L1 British English speakers familiar with Dutch English;
BrE-NF = L1 British English speakers not familiar with Dutch English;
L1-BrE = native speakers of British English; L1-AmE = native speakers of American English;
L1-NL = native speakers of Dutch; L1-Ger = native speakers of German;
L1-Spa = native speakers of Spanish; L1-Sing = native speakers of Singaporean English.

*** Method: VG = verbal guises; MG = matched guises.

moderate Dutch-accented Englishes compared to standard British accented English in terms of *speech understandability (intelligibility, comprehensibility, interpretability)* and *speaker evaluations (status, affect)* (Table 1). Furthermore, it was investigated whether the degree of Dutch-accentedness and familiarity with Dutch-accented English affected responses.

Study 2 was conducted to investigate whether it is possible to create a reliable research method that allows for the development of representative matched-guise speech samples for studies that are aimed at comparing L1 and L2 English accents (Table 1).

The aim of study 3 was to understand how highly educated L1 Dutch speakers understand and perceive Dutch-accented English, compared to L1 English accents (standard British and American), in three professional communication contexts (Table 1). This study was inspired by the intense debates held in the Netherlands on the suitability of Dutch-accented English in specific communication contexts. Studying Dutch listeners' responses to Dutch-accented English was thought to provide insights into the language norms held by L2 English speakers, and to contribute to our knowledge of the impact of accents on speakers by studying responses to accents in different communication contexts.

Study 4 investigated the responses of L2 English (German, Singaporean, Spanish) listeners to Dutch-accented English versus standard British and American English accents. The study replicated study 3 in terms of the employed research variables and technique, but used the three abovementioned L2 English speakers as listeners (Table 1), which allowed us to not only assess the responses to L2 and L1 English accents by various L2 English speakers, but also to understand whether the status of English as a global lingua franca has resulted in L2 English speakers that are able to flexibly deal with L2 English accents and what this might mean for perspectives on second language acquisition, LFE speech communities, accentedness research, and English language education.

Chapters 2-5 present the publications of the four studies. As this is a thesis based on publications, the content of the individual articles/chapters may overlap in places. Chapter 6 summarizes the main findings from the four studies, interprets them from the perspective of English as a global language and lingua franca, offers recommendations for language learning and teaching, and discusses implications for future research.

1.3 STUDY 1: BRITISH LISTENERS' RESPONSES TO DUTCH-ACCENTED ENGLISH

As stated in section 1.1, in the Netherlands, learners of English are taught L1 English varieties, mostly standard British English with an increased use of standard American English. Traditionally, the best learner of English is regarded as the learner who has become as native as possible, also in terms of pronunciation. However, it is not clear whether actually achieving a more native-like English accent is more advantageous to Dutch speakers of English compared to possessing a Dutch accent in English, for example in communications with L1 English speakers. In order to understand whether two degrees of Dutch-accented English and a standard British English accent evoked different reactions in native listeners, their responses in terms of *speech understandability* and *speaker evaluations* were assessed in the first study. Three elements of *speech understandability* were incorporated, based on Kachru and Smith (2008): intelligibility (distinguishing individual words and phrases), comprehensibility (understanding the meaning and intentions of an utterance), and interpretability (understanding the purpose of a communicative act). *Speaker evaluations* were assessed by means of two commonly studied speech evaluation research dimensions (Brown, 1965; Giles & Powesland, 1975; Ryan & Giles, 1992; Zahn & Hopper, 1985): status (the perceived competencies and social status of a person), and affect (the solidarity felt towards or a person or their likeability). The potential correlations between *speech understandability* and *speaker evaluations* were also investigated in order to understand whether there might be a connection between a listener's understanding of someone's speech and their evaluations of the speaker. A verbal-guise experiment was conducted with highly educated British professionals reacting to standard British English and two degrees of Dutch-accented English (slight and moderate). The three accents were produced by two speakers per accent in the context of a telephone sales talk. The British listeners were divided into listeners who were familiar with Dutch-accented English, and those who were not, because it was assumed that familiarity with an accent, both L1 and L2 accents, might impact the perceptions listeners have of a speaker and might influence speech understanding (Varonis & Gass, 1982; Fayer & Krasinski, 1987; Major et al., 2005; Smith & Nelson, 2006; Wang, 2007).

1.4 STUDY 2: REFINEMENT OF THE MATCHED-GUISE TECHNIQUE FOR THE STUDY OF THE EFFECT OF L2 ACCENTS COMPARED TO L1 ACCENTS

The first study described in section 1.3 employed a verbal-guise technique to measure reactions to accents. This is a commonly used research technique and allows researchers to collect speech samples or 'verbal guises' that are produced by L1 speakers of a language (variety). Each speaker produces only their own L1 language and no other language (variety). Another commonly used research method in studies on reactions to speech and accents is the matched-guise technique. This technique differs from the verbal-guise technique in that one speaker produces all languages and/or language varieties tested. Thus it ensures that listeners respond only to the language variety and/or accent under study, and not to individual speech characteristics of a speaker. In research aimed at comparing responses to L2 English accents with other L2 and L1 English accents, the use of the matched-guise technique is rare (see Mai & Hoffmann, 2014), probably due to the fact that an L1 speaker of a language cannot simultaneously be an L2 speaker of that same language. Therefore, it is very challenging to find suitable matched-guise speakers who can produce accents that represent particular language or accent varieties, that is, varieties that are accepted by listeners as authentic and not performed or 'put on' (Guy & Cutler, 2011: p.139).

In order to develop and test an evidence-based research method and produce representative matched guises for this thesis, the second study was conducted. Non-linguists and non-language specialists, or 'naïve listeners', who were either L1 speakers of British English, American English, or Dutch were asked to evaluate the representativeness of matched guises for standard British English, standard American English, and typical Dutch-accented English. Listeners were asked to evaluate the degree of nativeness and standardness of the accents. For the L2, Dutch English accent, rather than having listeners judge the degree of standardness of a matched guise, they were asked to assess the degree to which a matched guise sounded typical for an L1 speaker of Dutch speaking English. Dutch English is not an official English variety therefore has no established standardized grammar, vocabulary or accent, and as a result asking about the standardness of Dutch-accented English could have confused listeners.

This is to the best of our knowledge the first known attempt to try and establish a research method that allows researchers to determine whether listeners indeed have, as has been found before, clear perceptions of what sounds 'native' or 'standard' or not (e.g. Grondelaers & van Hout, 2015; Smakman, 2006).

1.5 STUDY 3: DUTCH LISTENERS' UNDERSTANDING AND EVALUTIONS OF DUTCH, BRITISH AND AMERICAN ENGLISH ACCENTS IN THREE COMMUNICATION CONTEXTS

As noted in section 1.1, English is commonly used in various domains of Dutch society and whether this is desirable is the subject of frequent public debate. Interestingly, these debates are not only focused on communications with other L2 or L1 English speakers, but also on English communications between L1 Dutch speakers. For example, some Dutch students in English streams of bachelor and master programmes have indicated that their lecturers' strongly Dutch-accented English at times hinders *speech understandability* and has resulted in less effective lectures (e.g. Huygen, 2017), which has been confirmed in research on perceptions of understanding (Hendriks, van Meurs & Hogervorst, 2016). It has also been observed that Dutch students in higher education who had been instructed in English got lower course grades than Dutch students who had been instructed in Dutch (de Vos, 2019). However, what is not clear yet is whether these reported effects of Dutch-accented English are reflections of decreased actual understanding (*speech understandability*) and/or a reflection of negative *speaker evaluations* of Dutch-accented English by Dutch listeners.

In order to understand Dutch listeners' responses to Dutch-accented English, compared to two L1 English accents, in terms of *speech understandability*, and *speaker evaluations*, a matched-guise experiment was conducted. *Speech understandability* was studied by applying the same three dimensions of the first study, however operationalized slightly differently. Also similar to the first study, the third study included *speaker evaluations* measurements for *status* and *affect*. In addition, a speaker's 'dynamism' was measured. Dynamism refers to someone's activity level and enthusiasm, and how the self-presentation of a speaker is perceived by other, which differs from perceptions of a speaker's *status* or *affect*, as confirmed by van der Haagen (1998) and Grondelaers, van Gent and van Hout (2015). This dimension is more rarely applied in accentedness studies compared to *status* and *affect*, and has proven to not always yield very strong correlations in their factorial loadings (e.g. van der Haagen, 1998). Nevertheless, *dynamism* was added to the third (and fourth study) because it could potentially yield interesting results that are relevant in professional communication contexts. The potential correlations between *speech understandability* and *speaker evaluations* were also calculated to understand whether there might be a connection between understanding someone's speech and the evaluations of speakers.

In addition, the reported negative effect of Dutch-accented English in a higher education, lecture context might not be observed in other contexts In order to understand whether Dutch listeners' responses to Dutch-accented English, compared to standard British and American English accents are generally applicable or whether they vary across professional communication contexts, these accents were studied in three different professional contexts: (1) a lecture; (2) an art gallery audio tour; (3) a job pitch. These contexts represent three domains in which English is commonly used as a lingua franca: higher education, tourism, and international business (Gerritsen et al., 2016; Edwards, 2016; Nickerson, 2005).

1.6 STUDY 4: WHERE DOES A 'FOREIGN' ACCENT MATTER? GERMAN, SPANISH, AND SINGAPOREAN LISTENERS' REACTIONS TO DUTCH-ACCENTED ENGLISH COMPARED TO STANDARD BRITISH AND AMERICAN ENGLISH ACCENTS IN THREE COMMUNICATION CONTEXTS

Since English is mostly used by L2 English speakers as a lingua franca to communicate globally, it is important for L2 English speakers to know if they are understood and how they are perceived by other L2 English speakers. In addition, understanding the responses to L2 English by L2 English speakers provides insights into the linguistic reality of English as a global language and lingua franca, and could inspire researchers and teachers of English as a foreign language to rethink what it means to be fluent or what constitutes successful language production and use. Traditionally, second language acquisition is considered most successful when language learners obtain proficiency in a target language free from any influence from their own L1. Some have resisted this premise and believe that English as a global lingua franca is a good example of a new, heterogeneous linguistic reality that has led to an international group of L2 English speakers or a speech community to arise, not separated by traditional national and linguistic boundaries (e.g. Canagarahah, 2007; Firth & Wagner, 1997). In this speech community, English is a tool to achieve professional or personal objectives, and as a result, speech community members and speakers of L2 English might have flexible attitudes towards other L2 English language use and is it no longer relevant to become as native as possible.

In order to further understand whether L2 English speakers indeed have adopted flexible attitudes towards L2 English language use and whether these attitudes can be observed in a study on the responses to Dutch-accented English, study 4 was

conducted. The final study assessed L2 English listeners' reactions to Dutch-accented English compared with standard British and American English in terms of *speech understandability* (intelligibility, comprehensibility, interpretability) and *speaker evaluations (status, affect, dynamism)* in three communications contexts (lecture, audio tour, job pitch). The L2 English listeners selected represent important international partners to the Netherlands and listeners with various linguistic backgrounds: Germany, Spain, and Singapore. Furthermore, the listener groups vary in geographical proximity to the Netherlands as well which might affect their general familiarity with Dutch-accented English. To investigate the effects of listener group, accent and context on speech understandability and speaker evaluations, a matched-guise study was conducted, with the same speakers from study 3,in which we compared three listener groups' (Germany, Spain, Singapore) responses to three English accents (Dutch-accented English, standard British and American English) in three communication contexts: a marketing lecture; an art gallery audio tour; a sales manager job pitch.

The following four chapters (Chapters 2-5) consist of the four studies that were conducted as part of this thesis. The first two studies were published and the final two studies have been accepted for publication with revisions.

2
RESPONSES TO DUTCH-ACCENTED ENGLISH

ABSTRACT

This paper reports on a study into the reactions of native speakers of British English to Dutch-English pronunciations in the onset of a telephone sales talk. In an experiment 144 highly educated British professionals who were either familiar or not familiar with Dutch-accented English responded to a slight Dutch English accent, a moderate Dutch English accent or a standard British English (BrE) accent. These accents were rated on the personality traits status and affect, on their intelligibility (orthographic transcription), comprehensibility (identification of key words), and interpretability (paraphrasing the purpose of the message). Although BrE was more intelligible and comprehensible than both Dutch English accents, all three accents were equally interpretable. The results indicated that British English pronunciation evoked more status than both Dutch English accents, and both British English and the slight Dutch English accent commanded more affect than the moderate Dutch English accent.

This chapter is based on:
Nejjari, W., Gerritsen, M., Haagen. M. van der, & Korzilius, H. (2012). Responses to Dutch-accented English. World Englishes, 31 (2), 248-268.

2.1 INTRODUCTION

In 2005, more than one third of the citizens of the European Union who do not have English as a first language claimed that they know English well enough to hold a conversation in English, and in the Netherlands as many as 87% of the respondents made this claim (European Commission, 2006). This high percentage provides insight into the important role English plays in the daily lives of Dutch people. For example, English is a compulsory language in Dutch education, and is taught in the later years of primary school and all through secondary school. A growing number of secondary school pupils receive up to 50% of their education in English, and more and more university students receive their education almost completely in English. In general, the variety of English that is taught is 'standard British English', which is based on grammatical criteria determined in Departments of English at Dutch universities. Furthermore, English (especially American English) is heard on TV in sitcoms, soap operas, drama series, movies and music videos every day, and it is also increasingly used in advertisements, especially in those targeted at young people and children (Gerritsen, Korzilius, van Meurs & Gijsbers, 2000; Gerritsen, Nickerson, van den Brandt, Crijns, Dominguez & van Meurs, 2007). Dutch companies with stock market quotations all publish their annual reports in English, and a third of these organisations do not even publish a Dutch version (De Groot, 2008). In short, the use of English in the daily lives of Dutch people is omnipresent (cf. Gerritsen & Nickerson, 2004).

Since a native (British or American) accent is notoriously difficult to learn because of transfer from the mother tongue (Vermeulen & Kellerman, 1998), the Dutch have their own pronunciation of English, obviously consisting of Dutch phonological features with a combination of British and American features Gussenhoven and Broeders (1997), van der Haagen (1998) and van den Doel (2006) give extensive overviews of features of Dutch-accented English, the most salient features being r-colouring, devoicing of all final obstruents, Th-stopping, and the lack of an /e ~ æ/ contrast.

Traditionally, 'non-native' speakers of English who use English as an international language are encouraged to learn and use one of the standard forms of 'native' English, because a standard 'native' form is supposed to help communication in international contexts. In this context 'native' English refers to the English spoken in countries where it is the mother tongue of the majority of speakers (e.g. Britain, the United States, Australia). Yet despite the significant influence this view has had on education systems all over the world (including the Dutch education system), many

authors currently advocate more tolerance towards 'non-native' varieties or new varieties of English (e.g. Alexander, 1999; Jenkins, 2006; Kachru, 1983; Phillipson, 1992; Pennycook, 1998; Seidelhofer, Breiteneder & Pitzl, 2006; van Oostendorp, 2002). These new Englishes should be accepted because 'it is now more important to be able to talk "to" native speakers of English, and not "as" native speakers of English' (Alexander, 1999: p. 27). Alexander proposes a *core area of intelligibility* or understandability for 'non-native' speakers, and urges listeners to adopt a more flexible attitude towards 'non-native' varieties of English. However, in reality 'native' speakers may not have such a flexible attitude towards L2 and FL1 speakers of English, and some 'non-native' varieties of English are often felt to be less intelligible, comprehensible and interpretable than L1 varieties.

2.1.1 Attitudes toward non-native Englishes

Many have observed that having an accent has an effect on the **attitudes**[2] listeners have towards speakers of native and non-native varieties of languages. Cargile and Giles (1997) show that a speaker's accent influences native listeners' attitudes towards that speaker; listeners report more feelings of pleasure (affect) and feel more positive when listening to an L1 speaker than when listening to an L2 or an FL speaker. Their study further shows that it is the speakers' accent per se, and not the strength of their accent that influences listeners' reported feelings of pleasure. Other studies, too (see Munro & Derwing, 1995 a, b for an overview), show that L1 speakers transfer negative attitudes towards a foreign accent into negative attitudes towards speakers of the accent. Kalin, Rayko and Love (1980) demonstrate, for example, that a negative attitude towards a speaker's accent influences the judgment of the status of that speaker with respect to the perceived suitability for higher end job positions. L1 speakers are judged to be more suitable for those job positions than L2 and FL speakers. Lindemann (2002) shows that a negative attitude also has an impact on the communication strategies used by native speakers, in that they tend to interrupt L2 and FL speakers more often than fellow L1 speakers, and frequently pretend not to understand L2 and FL speakers. Finally, Phiko (1997) shows that learners of English,

1 L2 English speakers were divided into L2 and FL speakers of English due to publication requirements. The other chapters in this thesis only distinguish between L1 and L2 English speakers, as discussed in the Introduction of this thesis.

2 Attitudes as defined in this chapter refers to what is defined as *speaker evaluations* in this thesis.

too, have negative attitudes towards non-native English. These learners considered national standard varieties of English as being 'real English' and non-native varieties as being *'strange English'*.

2.1.2 Intelligibility, comprehensibility and interpretability

Successful communication between 'native' and 'non-native' speakers of English greatly depends on the mutual understandability of one another's speech. Kachru and Smith (2008) describe the ability to understand language as consisting of three elements: *intelligibility*, *comprehensibility*, and *interpretability*.

Kachru and Smith define **intelligibility** as the ability to recognize "a word or another sentence-level element of an utterance" (2008: p. 61). A non-native pronunciation of these words or sentence-elements can greatly influence the overall intelligibility, as is demonstrated in the decision of the Indian call centre First Source, stationed in Mumbai, to open a branch in Belfast, Northern Ireland, employing only Northern Irish call-operators. They did this, because they had discovered that British customers found the Indian accent of their Indian employees unintelligible and sometimes became rather frustrated (De Volkskrant, 2006). However, experiments by Derwing and Munro (1997) in which 'native' speakers of English were asked to orthographically transcribe Cantonese-, Japanese-, Polish- and Spanish-accented English phrases showed that a strong L2 accent does not necessarily disrupt full intelligibility in terms of word and utterance recognition.

The second aspect in understanding communication is **comprehensibility**. This involves the ability to recognize both the meaning of words expressed and the intentions expressed by the speaker in the proper context, or as Kachru and Smith explain, 'the contextual meaning of the word in a socio-cultural setting as well as the illocutionary force of an utterance' (2008: p. 62). They suggest that comprehensibility can be measured by having an utterance paraphrased or by asking questions about its content. A common non-native speaker's misunderstanding of an English phrase is interpreting 'how are you' as an opportunity to discuss a person's well-being in great detail instead of as a simple greeting whose paraphrase would be 'hello.' Fayer and Krasinski (1987: 313), have shown that accentedness does affect comprehensibility in that L1 speakers believe FL speech is more difficult to understand than L1 speech.

The third element of the understandability of communication involves understanding the purpose of a communicative act, or the **interpretability** (Kachru & Smith, 2008). According to Kachru and Smith, interpretability involves cultural competencies and the ability to understand the discourse strategies people use.

A phrase like "it was nice meeting you" is successfully interpreted if the listener knows this as a signal of the end of a conversation, and not as a genuine remark. An example Kachru and Smith give is how communication breakdowns between Japanese and American business partners can be attributed to the differences in information structures used in Japanese and American culture. A Japanese person answering a question will often give an explanation of the answer before giving the actual answer, and in an interaction with a Japanese person it is very important to listen carefully to what is not being said. This can confuse Americans who are perhaps not familiar with this discourse strategy, even though both parties are quite fluent in the language they communicate in.

2.1.3 Familiarity

One of the factors that may influence the attitudes toward an accent, its intelligibility, comprehensibility and interpretability is **familiarity** with a variety or more specifically with an accent. Several studies have shown that familiarity with an accent, i.e. knowing that accent and having been exposed to it for some considerable time, aids intelligibility, comprehension and interpretability (Fayer & Krasinski, 1987; Major, Fitzmaurice, Bunta & Balasubramanian, 2005; Smith as cited in Kachru 2008; Varonis & Gass, 1982; Wang, 2007). Major et al. (2005) show that a sizable number of studies demonstrate that unfamiliar accents, both 'native' and 'non-native' alike, are more difficult to comprehend than familiar accents, but that this unfamiliarity disadvantage is not clearly related to the degree to which languages differ from each other or the 'linguistic distance' between the varieties in question. This suggests that it is the exposure to the variety and not necessarily the individual characteristics of the variety that creates the advantage for the listener. However, other studies have shown that varieties that share similar features are better comprehended (e.g. Bisazza, 1982; Flowerdew, 1994; Wilcox, 1978; Bent & Bradlow, 2003; Major et al., 2005; Wang & van Heuven, 2007). Furthermore, Smith has shown that familiarity can aid interpretability (as cited in Kachru 2008), and more recently Wang (2007), too, observed that familiarity with an accent facilitates intelligibility. Moreover, there are indications that attitude, intelligibility, comprehensibility and interpretability influence each other (Smith, 1992; Lindemann, 2002), and that a positive attitude towards an accent, irrespective of familiarity, increases intelligibility and hence comprehensibility (Fayer & Krasinski, 1987).

The studies mentioned above clearly show that L1 speakers' attitudes tend to be more negative towards L2 and FL speakers of English than towards fellow L1 speakers. Whether FL and L2 English is really less intelligible, comprehensible and

interpretable than L1 English has not yet been fully established due to the limited number of studies and conflicting results. These conflicting results may, among other things, be due to differences in methodology and definitions used to measure intelligibility, comprehensibility and interpretability. For example, some methods test perceived comprehensibility (defined as the estimation of the respondents to indicate the extent to which the message is understood), while other methods use real-life interactions between L1 and FL speakers of English and include an evaluation of the success of the interaction (Lindemann, 2002). Moreover, respondents reporting their own assessment of comprehensibility and intelligibility may not necessarily reflect reality. Finally, only a small number of varieties of FL and L2 English have been studied, and it is plausible that other varieties from other sociolinguistic contexts will evoke different responses.

As we have shown in the beginning of this introduction English is taught to almost all citizens of the Netherlands and is used widely, especially in professional contexts. In order to improve the intercultural encounters in English between speakers of Dutch English and speakers of other varieties of English it is important to gain insight into the effect of Dutch English on other speakers of English and to find out whether improving the Dutch English accent improves the intercultural communication.

2.1.4 Research questions

The main purpose of this study is to investigate the reactions to Dutch-accented English pronunciations compared to a 'standard British English' pronunciation (BrE), and more specifically, to see whether different degrees of Dutch accentedness have different effects on the attitudes of 'native' speakers towards these accents, and on the intelligibility, comprehensibility and interpretability of these accents, and whether familiarity with Dutch-accented English plays a role in all of this. Research into the effect of Dutch-accented English on 'native' speakers of English is not entirely new. Koster and Koet (1993) and Koet (2007) investigated whether 'native' speakers of English and Dutch teachers of English had different responses to Dutch speakers speaking English, in terms of their pronunciation, and on how beautiful, monotonous, cultured and pleasant they found the accents. They found, rather unexpectedly, that the 'native' speakers of English were more tolerant than the Dutch teachers. Van den Doel (2006) researched which features of a Dutch accent bothered speakers of British and American English most. He came to a hierarchy of mistakes, with problems with word-stress on top, and which for the rest turned out to be different for 'standard British English speakers and for speakers of

'General American'[3] (van den Doel, 2006: 292). He suggests that these pronunciation features be given priority in education, and urges teachers to focus on the most problematic areas.

The present study is different from the studies just discussed, in that we are not only interested in the attitudes of British speakers to Dutch-accented speech, but that we compare these with their attitudes towards 'standard British English' (BrE). Furthermore we will not only look at the attitudes toward the accents (Dutch and 'standard British English'), but also at the intelligibility, comprehensibility, and interpretability of these accents. We will study the responses of 'native' speakers of British English to two levels of educated Dutch-accented English pronunciations: a moderate accent (Educated speakers) and a slight accent (trained International Business Communication students; for details, see the Method section). We selected these Dutch speaker groups, because they represent the speaker groups most likely to use English in a professional context. We will focus on the following three research questions.

- RQ1: To what extent do native speakers of British respond differently to a 'standard BrE' accent, a slight Dutch accent and a moderate Dutch accent in terms of:
  1A. The attitudes towards the speakers;
  1B. The degree of intelligibility;
  1C. The degree of comprehensibility;
  1D. The degree of interpretability?
- RQ2. What is the relationship between the attitude towards an accent, the degree of intelligibility, the degree of comprehensibility, and the degree of interpretability?
- RQ3. To what extent does familiarity with Dutch-accented English have an effect on the attitude of native speakers of British English towards these accents and on the rate of intelligibility, comprehensibility, and interpretability?

3 General American refers to what is defined as standard American English in this thesis.

2.2 METHOD

2.2.1 Respondents

The respondents were 72 'native' speakers of British English living in England in the London area, and 72 'native' speakers of British English who had been living in the Netherlands for ten years or longer. The latter were, of course, extremely familiar with Dutch-accented English, in the sense that they had been exposed to it for a considerable amount of time, while the former were not very familiar with Dutch-accented English. All respondents were, highly educated British citizens and successful international professionals; 30% of the respondents were male and 70% female, and 71% was aged 35 or older. According to the interviewer, who interviewed two-thirds of the respondents in a face-to-face situation and spoke to the others over the telephone (cf. 2.2.5 Procedure) all were speakers of 'standard British English' – though not necessarily Received Pronunciation (RP) – according to grammatical criteria used for the determination of 'standard British English' in Departments of English at Dutch universities.

2.2.2 Materials

As stimulus we used the onset of a telephone conversation, so that there were no non-verbal cues that might affect attitude, intelligibility, comprehensibility and interpretability (cf. Rubin, 1992). The message was an imitation of a telephone sales talk for a Dutch asset management business that wanted to expand their market in the United Kingdom (see section 2.5.1). There were six stimuli produced by six women. Two speakers were 'native' standard BrE near-RP speakers, two were (phonetically) highly trained MA students in the English stream of the International Business Communication programme at the Radboud University Nijmegen (IBC English, slightly accented). This level of English represents one of the highest levels of English Dutch people can obtain in the Netherlands and the degree programme aims at preparing students for an international business career. The two other Dutch-accented speakers were linguistically naïve university students from Nijmegen whose English was representative of Dutch people with the highest level of English one can acquire in Dutch secondary schools (moderately accented English). Thus the samples represented three degrees of accentedness: (1) BrE, (2) slightly Dutch-accented and (3) moderately Dutch-accented.

The samples were judged by a trained phonetician experienced in teaching English pronunciation to Dutch learners, and by an applied linguist who is a 'native' speaker of English and who is accustomed to degrees of Dutch-accentedness. In addition, 50

undergraduate students of English Language and Literature evaluated the samples. Both expert and non-expert judges confirmed that all samples were representative of their levels of accentedness and possess the typical pronunciation characteristics of Dutch-accented English and 'standard British English' (Gussenhoven & Broeders 1997; van den Doel, 2006). In order to see if we could simply take the means for the two speakers in each of the groups as representative for the three degrees of accentedness, we performed a number of statistical tests. One-way ANOVAs were performed on variables of interval level, and when the variables were of a nominal or ordinal level, cross table analyses were done using Chi-square tests. One-way ANOVAs and Chi-square tests revealed that there were no statistically significant differences between the two speakers within each speaker group on any of the dependent variables studied. Due to the large quantity of statistical information (Ms and SDs and contingency tables) results for the three speaker groups are withheld.

2.2.3 Design

The design of the study was a 3 x 2 between subject factorial design (Table 1), where each of the 144 respondents was randomly but evenly assigned one of the six speech samples, and was asked to perform four tasks that measured:

- the attitude towards the speaker;
- the intelligibility of the speech sample;
- the comprehensibility of the speech sample;
- the interpretability of the speech sample.

Thus, each sample was judged by 24 listeners. A power analysis reveals that with these numbers, we expect to find large differences between the groups (a large effect size, f=.40) with the statistical tests used at an alpha level of .05 for 99% of the cases (statistical power=.99) (see Cohen, 1988).

Table 1 Design experiment (N=144).

Listeners / Speakers	Listeners familiar with Dutch-accented English (n=72)	Listeners not familiar with Dutch-accented English (n=72)
Standard BrE Accent	24	24
Slightly Accented	24	24
Moderately Accented	24	24

2.2.4 Measuring instruments

For the attitude task (RQ1A), each respondent listened to the speech sample of one speaker and rated that speaker on eight five-point semantic differentials based on previous studies (van der Haagen, 1998; Gerritsen et al., 2000): *competent-incompetent, irritating-pleasant, educated-uneducated, aggressive-considerate, intelligent-stupid, inferior position-having authority, cultured-not cultured and unfriendly-friendly*. The answers were, where necessary, recoded so that all the scales ranged from 1 as the most negative rating and 5 as the most positive rating. A factor analysis, using a principal axis factoring extraction method with an Eigenvalue > 1 criterion for factor extraction, followed by a varimax rotation, on the ratings of these semantic differentials showed a resolution into two factors.

In language attitude research we frequently find that there are two or three dimensions operative in evaluations. Brown (1965) claims there are two norms that determine social interaction, namely the status norm and the solidarity norm, and many researchers have since used these terms to label the factors that play a role in evaluating accents (e.g. Giles & Powsland, 1975; Ryan & Giles, 1992). However, most attitude research has been done on L1 speakers rating accents or dialects from that same L1, and in that case it makes sense to speak about the "solidarity" a listener feels with the (accent of) the speaker. In our case, where L1 listeners rate L2 speakers, and where the listeners were highly educated business professionals and the speakers young university students, solidarity would not be a sentiment one would expect the listener to feel. For this reason we have decided to follow Van der Haagen (1998), and label the second dimension affect.

Table 2 shows the factor loadings, i.e. the correlation coefficients, of the items with these two factors. The items are grouped such that the first five correlate highest with factor one and the next three correlate highest with factor two. The scales contributing to Factor 1 are *competent*, *educated, having authority*, *intelligent*, and *cultured*, which suggests that this factor is a measure of the status the subjects attributed to the speakers; the scales contributing to Factor 2 are *considerate*, *pleasant* and *friendly*, so that this factor can be seen as a measure of (personal) *affect* the subjects have for the speakers. The reliability, in terms of Cronbach's alpha, of both factors was adequate (>.70). For each factor composite scores were calculated by taking the mean of the ratings on the scales contributing to that factor.

For the **intelligibility** task (RQ1B), the respondents were asked to listen to the first two sentences from the sample again (see section 2.5.2), and like in the study by Munro, Derwing and Morton (2006) and by Kachru and Smith (2008) respondents

Table 2. Rotated Factor Matrix: factor loadings of the scores on nine scales with two factors. Only loadings >.30 have been printed.

	Factor 1	Factor 2
Competent	.83	
Educated	.80	
Having authority	.78	
Intelligent	.77	
Cultured	.72	
Considerate		.86
Pleasant	.48	.62
Friendly		.76
% variance explained	39	23
Reliability Cronbach's α	.85	.74

were asked to transcribe these two sentences orthographically. The respondents were allowed to listen to these two sentences no more than two times. Intelligibility was measured by counting the number of wrongly transcribed words, a method described by Munro, Derwing and Morton (2006). A speaker was considered completely intelligible if all the thirty-three words were transcribed correctly. Each word that was either misspelled or replaced by another word was counted as one error. The only error in the transcription that was not counted was SNG instead of SMG, because of the extreme similarities between the nasal consonants /n/ and /m/ (Gussenhoven and Broeders 1997).

In order to measure **comprehensibility** (RQ1C) and **interpretability** (RQ1D), we used Kachru and Smith's definitions of these variables. For **comprehensibility** we asked the respondents to paraphrase the first two sentences of the sales pitch (see section 2.5.2). We decided that the key words, for these sentences were (1) *Dutch*, (2) *asset*, (3) *management*, (4) *expand*, (5) *market*, and (6) *Great Britain*, and we counted for each respondent how many keywords they had written down. In order to measure interpretability we asked the respondents to write down the purpose of the message of the first two sentences of the sample (see section 2.5.2). The following three answers to the question about the purpose of the recorded text were considered correct: (1) *a business pitch*, (2) *an attempt to sell a product/service*, or (3) *a cold call to attract potential clients/business partners*. A message was considered interpreted incorrectly if a respondent did not write that the sample was either a business pitch,

an attempt to sell a product/service, or a cold call to attract potential clients/business partners.

2.2.5 Procedure

Two-thirds of the respondents performed the tasks in a one-on-one situation with the first author, either face-to-face or over the telephone, and one-third did them on their own via the Internet. The speech samples were equally distributed among the three test situations. An independent t-test showed that the correlation between the test condition and the attitudes,, intelligibility, comprehensibility was not significant (p's > .05). A Chi-square test showed that there were no significant correlations between the test condition and interpretability ($p > .05$).

2.2.6 Statistical analyses

All analyses were conducted with SPSS 15.0. For the attitude scales where Cronbach's α was at least adequate (.70 or higher) we calculated composite means of the items, which was the case for all scales. Several One-way ANOVAs were performed with the independent variables accentedness (British English, slightly Dutch-accented English, moderately Dutch-accented English) and familiarity with Dutch-accented English (familiar, not familiar), and using attitudes (status, affect), intelligibility, comprehensibility, and interpretability as dependent variables. If the differences were significant, a post-hoc Bonferroni procedure was used when variances were equal, and a post-hoc Games-Howell when variances were unequal (tested with Levene's test for equality of variances). To measure the differences between the speaker groups in the identification of the key words for the comprehensibility task, a Mann-Whitney was calculated. A Mann-Whitney test is the non-parametric alternative of the t-test for independent samples and looks at differences in the ranked positions of ordinal dependent variables in two independent groups. Mann-Whitney gives the probability that the outcome is a chance result testing the null-hypothesis that the two groups are equivalent in rank positions. Since the dependent variable interpretability was nominal, a Chi-square test was performed to measure the differences for interpretability between the speaker groups. The effects of familiarity with Dutch-accented English for the two Dutch accents (slightly and moderately Dutch-accented) on the dependent variables attitudes, intelligibility and comprehensibility were studied by means of Two-way and One-way ANOVAs. For the effects of familiarity on the interpretability of the accents, a Chi-square test was used. The relationship between all four dependent variables was investigated using Pearson Correlations tests.

2.3 RESULTS

2.3.1 Attitudes towards Dutch-accented English (RQ1A)

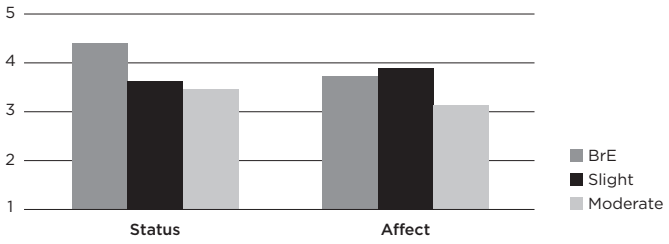
This section will present the results for *status* and *affect* for the three speaker groups for all the listeners together (familiar and not familiar). The results for both the factors *status* and *affect* are presented in Table 3, where we see the mean ratings and standard deviations for the three accents, and the results from One-way ANOVAs with the significant differences between the accents according to post-hoc contrasts.

Table 3. Mean STATUS and AFFECT per accent for all listeners (N=144).

	M	SD	Differences between speaker groups
STATUS			
BrE (n=48)	4.44	0.48	
Slight (n=48)	3.65	0.62	BrE vs Slight***and BrE vs Moderate***
Moderate (n=48)	3.50	0.62	
AFFECT			
BrE (n=48)	3.75	0.83	
Slight (n=48)	3.90	0.74	Slight vs Moderate*** and BrE vs Moderate***
Moderate (n=48)	3.15	0.52	

Notes: Mean scores (M), standard deviations (SD),
*** = p < .001; 1=negative, 5=positive; n=48 per accent; N=144 all listeners.

Figure 1. Mean scores STATUS and AFFECT
(1=negative, 5=positive) per accent (n=48) for all listeners (N=144).



Status

The results for status are shown in Table 3 and Figure 1. A One-way ANOVA revealed that the differences in perceived status for the three accents were significant ($F(2, 141)=34.74$, $p<.001$; $\eta^2=.18$). Post-hoc Bonferroni tests showed that the BrE speakers were considered to have significantly more status than both the slightly accented Dutch English speakers and the moderately accented Dutch English speakers (p's < .001).

Affect

Table 3 and Figure 1 show the results for *affect*. A One-way ANOVA revealed that the differences in *affect* for the three accents were significant ($F(2, 141)=15.12$, $p<.001$; $\eta^2=.33$). A post hoc Games-Howell test showed that the BrE speakers and the slightly accented speakers did not differ significantly from each other in this respect and that both groups commanded significantly more *affect* than the moderately accented speakers (p's < .001).

2.3.2 Intelligibility of Dutch-accented English (RQ1B)

The results for *intelligibility* are presented in Table 4, where we see the mean ratings and standard deviations for the three accents and the results from a One-way ANOVA with the significant differences between the accents according to post-hoc contrasts. In general, the intelligibility of all speaker groups was high; all respondents correctly transcribed at least 29 out of 33 words. The 'standard BrE' speakers turned out to be the most intelligible speakers, followed by the moderately accented speakers and the slightly accented speakers. However, as is shown in Table 4, the standard deviations suggest that the transcriptions for the 'standard BrE accent' were more homogeneous than for the other two accents. The differences in the intelligibility of the three speaker groups were significant ($F(2,140)=12.65$, $p<.001$; $\eta^2=.15$). A post-hoc Games-Howell test revealed that both Dutch accented groups were less intelligible than the BrE speaker group (p's < .01).

Table 4. Mean Intelligibility per accent for all listeners (N=144).

	M	SD	Differences between speaker groups
BrE (n=48)	32.67	0.60	
Slight (n=48)	29.92	3.71	BrE vs Slight*** and BrE vs Moderate***
Moderate (n=47)	30.40	3.23	

Notes: Mean scores (M), standard deviations (SD),
*** = p<.01; min=0 words, max=33 words.

2.3.3 Comprehensibility of Dutch-accented English (RQ1C)

In order to investigate the extent to which British native speakers responded differently to a 'standard BrE accent', a slight Dutch accent and a moderate Dutch accent in terms of the degree of comprehensibility, a One-way ANOVA and a Mann-Whitney were executed. All respondents were able to identify at least five out of the six key words (see Tables 5 and 6). The BrE speaker group was most comprehensible (M=5.94), followed by the slightly accented speaker group (M=5.17) and the moderately accented speaker group (M=5.00). A One-way ANOVA showed that there was a significant difference between the speaker groups in the degree of comprehensibility ($F(2,141)=12.88$, $p < .001$; $\eta^2=.15$). Post-hoc Games-Howell tests revealed that both Dutch accented groups were less comprehensible than the BrE speakers (p's < .01).

The identification of the key words for the comprehensibility tasks showed interesting patterns since certain words were identified more often for specific speaker groups than others. Table 6 shows the number of respondents identifying the six key words per speaker group. We see that for almost all groups the words 'expand' and 'Great Britain' were mentioned, but that for the other four words there is

Table 5. Mean *Comprehensibility* per accent for all listeners (N=144).

	M	SD	Differences between speaker groups
BrE (n=48)	5.94	0.25	
Slight (n=48)	5.17	1.12	BrE vs Slight *** and BrE vs Moderate***
Moderate (n=48)	5.00	1.22	

Notes: Mean scores (M), standard deviations (SD),
*** = p < .01; min=0, max=6; n=48 per accent; N=144 all listeners.

Table 6. Frequencies identified key words per accent for comprehensibility for all listeners (N=144).

	Dutch	Asset	Management	Expand	Market	Great Britain
BrE (n=48)	48	47	48	48	46	48
Slight (n=48)	41	42	40	45	33	46
Moderate (n=48)	42	32	36	46	40	46

Notes: n=48 per accent; N=144 all listeners.

variability between the speaker groups. These differences were calculated by means of a Mann-Whitney test. For the British English speakers the following words were identified significantly more often than for the slightly accented speakers: 'Dutch' ($Z=2.73$; $p<.01$), 'management' ($Z=2.94$; $p<.01$) and 'market' ($Z=3.46$; $p<.01$).

For the British English speakers the following words were identified more often than for the moderately accented speakers: 'Dutch' ($Z=2.52$; $p<.01$) 'asset' ($Z=3.99$; $p<.001$), 'management' ($Z=3.68$; $p<.001$), and 'market' ($Z=1.99$; $p<.05$). For the slightly accented group the word 'asset' was significantly more often identified than for the moderately accented group ($Z=2.42$; $p<.01$). Furthermore, the word 'asset' was identified as 'acid' nine times for the moderately accented group.

2.3.4 Interpretability of Dutch-accented English (RQ1D)

Table 7 shows for each accent how many listeners correctly interpreted the business pitch. A Chi-square test showed there were no significant differences in the interpretability of the three accents ($\chi^2(2)=.63$, $p=.73$).

Table 7. Interpretability per accent for all listeners (N=144).

	Correctly Interpreted	
	n	%
BrE (n=48)	42	87.5
Slight (n=48)	44	91.7
Moderate (n=48)	44	91.7

Notes: n=48 per accent; N=144 all listeners.

2.3.5 Relationship between attitudes, intelligibility, comprehensibility and interpretability (RQ2)

Pearson Correlations (Table 8) showed that there was a significant positive correlation between the factors *status* and *affect*, *status* and *intelligibility*, *status* and *comprehensibility*, and *affect* and *comprehensibility*, and *intelligibility* and *comprehensibility*. We realise that the scores for the factors attitude (status, affect), intelligibility and comprehensibility are scales, while those for interpretability are dichotomous, so that we should not treat their correlations uniformly. However, Point Biserial and Spearman tests yielded virtually identical results.

Table 8. Pearson Correlations for status and affect, intelligibility, comprehensibility, and interpretability for all listeners (N=144).

	Status	Affect	Intelligibility	Comprehensibility	Interpretability
Status	1				
Affect	0.37*	1			
Intelligibility	0.18**	0.11	1		
Comprehensibility	0.22**	0.17**	0.84*	1	
Interpretability	0.09	0.16	-0.15	-0.16	1

Notes: * $p < .001$; **$p < .05$.

2.3.6 Familiarity with Dutch-accented English and attitude, intelligibility, comprehensibility and interpretability (RQ3)

To investigate the effects of familiarity with Dutch-accented English (familiar, not familiar) on the attitudes towards both Dutch English accents and on the intelligibility and comprehensibility of Dutch English pronunciations, Two-way and One-way ANOVAs were conducted with the factors familiarity, accentedness, attitude, intelligibility and comprehensibility. The effect of familiarity with Dutch-accented English on the interpretability of Dutch English was measured using a Chi-square test.

Table 9 shows the mean ratings and the standard deviations of the listener groups familiar and not familiar with Dutch-accented English for the slight and moderate Dutch accents for attitude (*status*, *affect*), intelligibility and comprehensibility. A Two-way ANOVA showed no main effects for the two listener groups (familiar, not familiar) and accentedness (Slight, Moderate) with respect to the evoked *status* ($F(1, 96) = 2.30$, $p > .05$, $\eta^2 = .07$) and *affect* ($F(1, 96) = 2.61$, $p > .05$, $\eta^2 = .02$). Furthermore, a Two-way ANOVA also showed no main effects for the two listener groups (familiar, not familiar) and accentedness (Slight, Moderate) for the intelligibility ($F(1, 96) = 0.73$, $p > .05$, $\eta^2 = .05$) and comprehensibility ($F(1, 96) = 0.31$, $p > .05$, $\eta^2 = .11$).

A One-way ANOVA with the factor familiarity showed no significant differences between the listener groups for *affect* ($F(1, 94) = 1.66$, $p > .05$, $\eta^2 = .02$) and comprehensibility ($F(1, 94) = 2.52$, $p > .05$, $\eta^2 = .03$). Listeners familiar with Dutch-accented English, however, assigned to both the speakers of the moderate and the slight Dutch accent significantly less status ($M = 3.54$) compared with the listeners not familiar with Dutch-accented English ($M = 3.86$) ($F(1, 94) = 6.64$, $p > .05$, $\eta^2 = .07$), and found them more intelligible ($M = 31.08$) in comparison with listeners not familiar with Dutch-accented English ($M = 29.25$) ($F(1, 94) = 7.22$, $p < .01$, $\eta^2 = .07$).

Table 9. Familiarity per Dutch-accented English accent for attitude (1=negative, 5=positive), intelligibility (Min=0, Max=33), comprehensibility (Min=0, Max=6) for the 96 listeners who listened to a Dutch-accented English fragment.

		Number of listeners	M	SD
STATUS				
Both Dutch English accents	Familiar	48	3.54	0.56
	Not familiar	48	3.86	0.65
Slight	Familiar	24	3.63	0.56
	Not familiar	24	3.69	0.68
Moderate	Familiar	24	3.23	0.46
	Not familiar	24	3.76	0.65
AFFECT				
Both Dutch English accents	Familiar	48	3.42	0.77
	Not familiar	48	3.62	0.71
Slight	Familiar	24	3.90	0.66
	Not familiar	24	3.89	0.83
Moderate	Familiar	24	2.94	0.54
	Not familiar	24	3.35	0.43
Intelligibility				
Both Dutch English accents	Familiar	48	31.08	2.67
	Not familiar	48	29.25	3.90
Slight	Familiar	24	31.13	2.31
	Not familiar	24	28.71	4.44
Moderate	Familiar	24	31.04	3.04
	Not familiar	24	29.79	3.28
Comprehensibility				
Both Dutch English accents	Familiar	48	5.27	1.14
	Not familiar	48	4.90	1.17
Slight	Familiar	24	5.33	1.01
	Not familiar	24	5.00	1.22
Moderate	Familiar	24	5.21	1.29
	Not familiar	24	4.79	1.14

Notes: Mean scores (M), standard deviations (SD).

As is shown in Table 10, for *interpretability*, the listener groups familiar with Dutch-accented English correctly interpreted the intentions of the speaker in 90% of all cases, and the listeners not familiar with Dutch-accented English correctly interpreted the speaker's intentions in 94% of all cases. The difference in percentages for

the two listener groups is not statistically significant. The two listener groups also interpreted the messages with the slight accent equally well as the ones with the moderate accent.

Table 10. Familiarity (n=48) per Dutch English accent and Interpretability for the 96 listeners who listened to a Dutch English fragment.

	Correctly Interpreted	
	n	%
Familiar (n=48)	43	90
Slight (n=24)	21	88
Moderate (n=24)	22	92
Not familiar (n=48)	45	94
Slight (n=24)	23	96
Moderate (n=24)	22	92

2.4 CONCLUSION AND DISCUSSION

Our first research question investigated the extent to which British 'native' speakers respond differently to a 'standard British English accent', a slight Dutch accent and a moderate Dutch accent in terms of (RQ1A) attitudes towards the speakers (RQ1B) the degree of intelligibility; (RQ1C) the degree of comprehensibility, and (RQ1D) the degree of interpretability of the three accents.

2.4.1 Attitudes

For attitudes (RQ1A), the finding that all listeners attributed a higher status to the BrE speakers than to both groups of Dutch-accented speakers are in line with Munro and Derwing (1995 a, b). There, too, L1 speakers were attributed more *status* than FL speakers. These results are also in line with a recent study by Coupland and Bishop (2007: 79) in which the attitudes towards different 'native' and 'non-native' English accents were measured. They showed that in the UK 'standard English' evokes the highest social attractiveness and the second most prestige. Overall, the respondents felt equal *affect* for the BrE speakers and the slightly accented Dutch speakers, and they felt less *affect* for the moderately accented Dutch speakers. This suggests that listeners not only respond to accentedness itself, but also to degrees of accentedness. Interestingly, the slightly accented speakers command as much *affect* as the British

English speakers. Our findings suggest that for L1 listeners the accent should not be too strong either, since the moderately accented speakers command significantly less *affect* than the slightly accented speaker.

2.4.2 Intelligibility

To test the intelligibility of the three accents (RQ1B), the listeners were asked to transcribe two sentences. The British English speakers were more intelligible than the slightly accented and the moderately accented Dutch speakers. Similar to findings by Derwing and Munro (1997), we found that the non-native accents were at least reasonably intelligible. However, these results are not in line with their implicit assumption (shared by Varonis & Gass 1982) that a stronger accent would have a negative influence on the intelligibility, since both Dutch accents were equally intelligible. It appears that in the case of the Dutch English accents studied, the foreign accent itself, rather than the degree of accentedness decreases intelligibility.

2.4.3 Comprehensibility and interpretability

To test differences and similarities between the comprehensibility and interpretability of the three accents (RQ1C, 1D), the listeners were asked to write down what they thought the purpose of the message was. The 'standard British English accent' was more comprehensible than slightly and moderately Dutch-accented English. This result is similar to Fayer and Krasinski (1987), who claim that accent does affect comprehensibility in that L1 speakers understand messages less well when uttered by FL speakers. The identification of the key words for the comprehensibility task shows an interesting pattern in that British listeners were able to observe the two degrees of accentedness for the word 'asset', which was in nine cases identified for the moderate Dutch English accent as 'acid'. This confusion is understandable, since most naïve Dutch speakers of English are not aware of the phenomenon of pre-fortis clipping (Gussenhoven & Broeders 1997), i.e. the shortening of vowels before fortis obstruents. This means that they make the second vowel in 'asset' too long, which in its turn makes the word sound like 'acid' to British ears. So in this case the additional pronunciation training Business Communication students, the speakers of the slight Dutch English accent, receive seems to have resulted in a higher comprehension for the slight accent.

Finally, the results for interpretability show that L1 speakers understand the intentions of L1 and FL speakers equally well (Table 7). It needs to be noted, however, that all respondents were international professionals and would be familiar with the sales tactics used in this experiment, i.e. cold calling potential clients or business

partners and pitching products and/or services. The respondents' possible familiarity with specific business tactics, the cultural barrier that might not be as large as the example of American and Japanese culture given by Kachru and Smith (2008), and the high interpretability of the business pitch seem to confirm Smith's assertion that familiarity can aid interpretability (as cited in Kachru, 2008). However, that familiarity need not be with the language or accent of the FL speaker, but rather with the situational and social context.

2.4.4 Relationship between attitudes (status, affect), intelligibility, comprehensibility and interpretability

The second research question investigated the possible relationship between the attitudes towards an accent, its degree of intelligibility, comprehensibility and interpretability. The five positive correlations suggest that it is worth the effort for a speaker of English as an FL to use a variety that is associated with a higher status, since our study shows that a high status is related to more affect towards a speaker and a higher intelligibility and comprehensibility. In addition, using an intelligible variety of English will aid comprehensibility, which in turn leads to more affect towards a speaker. The correlation between status and intelligibility is consistent with similar findings by Fayer and Krasinski (1987), and Smith (1992).

2.4.5 Familiarity

The third research question asked whether familiarity with Dutch-accented English might have an effect on the attitudes of native speakers of British English towards these accents and on the intelligibility, comprehensibility and interpretability of the two Dutch accents. Our results show that listeners familiar with Dutch-accented English consistently rated the Dutch speakers significantly lower on *status* than the listeners who were relatively unfamiliar with the Dutch accent, and that those familiar also rated them lower on *affect*, albeit not significantly. Familiarity with Dutch-accented English resulting in a lower perceived status may be due to the fact that a stronger accent suggests a poorer fluency in English and is perhaps associated with a lower educational level and therefore with a lower social status. Another possible explanation could be that the listeners living in the Netherlands have much more contact with the Dutch and may have some negative feelings towards some of them. In any case, they *knew* they were rating Dutch FL speakers, whereas the listeners not familiar with a Dutch accent were rating "just a non-native accent", although of course the word Dutch was mentioned in the stimulus materials. Similar results have been found in a large scale project started in New Zealand and copied in

many countries world-wide (Bayard, Weatherall, Gallois & Pittam, 2001; Bayard et al., 2003) All in all, our conclusion is that as far as attitudes are concerned, it does not help from a speaker point of view if the interlocutor is familiar with an accent, because (s)he will assign a speaker less *status* and *affect*.

On the other hand, our results also show that those familiar with a Dutch accent find the Dutch speakers significantly more intelligible and not significantly but consistently more comprehensible and interpretable. If we take this, along with earlier findings (e.g. Fayer & Krasinski, 1987; Major et al., 2005; Varonis and Gass 1982; Bent & Bradlow, 2003; Wang & Van Heuven, 2007; Wang 2007) it seems we have to concur that it helps to avoid communication breakdowns if an interlocutor is familiar with an accent.

It would be interesting to see whether British expats living in other countries and Britons in the UK have similar attitudes towards other non-native English accents compared to standard British English or other native varieties of English. In short, the inclusion of the factor familiarity with a Dutch accent in this study of language attitudes has shown different responses to varieties of English pronunciations (native and non-native), and seems to be an interesting topic for further investigation.

2.4.6 Dutch English as an international business communication variety

The main purpose of our investigation was to see if the effects of FL English on native speakers described in the literature would also apply to Dutch-accented English, and if so, whether there would be different effects depending on the degree of accentedness. Our results show that indeed standard British English leads mostly to more positive attitudes and is more intelligible and comprehensible, but it is not easier to interpret than both Dutch accents. However, both degrees of Dutch accentedness actually are evaluated positively and are well-understood; for the factors *status* and *affect* the scores are never lower than 3 (where 1 is most negative, 3 is neutral, and 5 most positive), at least 91% of the speech is intelligible and at least 83% is comprehended. In short, the reception of Dutch-accented English is, in a telephone setting, not unfavourable. The question then arises whether we need to improve the English pronunciation of Dutch speakers.

If we look at the interpretability of both levels of Dutch-accented English we see that the speakers' intentions are as well understood as the British English speaker's intentions. If the aim of teaching English is only to make students become interpretable, improvement of the pronunciation of Dutch learners is unnecessary since their intentions are understood, even if some of the words are not intelligible or comprehensible. Apparently listeners do not need to hear every single word to

understand a text and our research shows that Alexander (1999) is right; native speakers will try to understand the message a FL speaker is trying to convey, and … they succeed. However, for successful communication not only the message itself, but also the impression a speaker makes on the listener plays a role, and in this respect the native speakers turn out not to be as flexible as Alexander assumes, since they attribute significantly less status to both levels of Dutch English than to British English.

This suggests that it might be worth the Dutch learners' while to try to achieve a standard British English accent. Furthermore, aiming for an accent that carries status can perhaps pay off, since *status* is positively correlated with intelligibility, (and intelligibility in its turn is positively correlated with comprehensibility) and with *affect*. Improving one's pronunciation will thus kill a number of positive birds with one stone.

One of the most remarkable results from our investigation is that the only differences between the two levels of Dutch-accentedness can be found in the attitudes of the listeners. For *affect*, speakers with a slight accent score better than speakers with a moderate accent. The question arises whether for speakers with a slight Dutch English accent it is worth the effort to get to the level of accentedness they achieve if reaching that level only pays off in evoking more favourable attitudes. In our view, it is worth the effort because the attitude towards a speaker can influence the listeners' behaviour.

2.4.7 Suggestions for further research

By using the onset of a telephone conversation as our setting rather than a face-to-face situation, we have eliminated all extra-linguistic cues that can play a role in communication and have measured responses to the accent *per se*. This has allowed us to gain insight into the responses of British native speakers to two degrees of Dutch-accented English, but we do realize that in communication in a different setting various other cues, such as non-verbal communication, grammar, lexis etc., obviously also play a role and should be investigated.

We have tried to gain insight into the intelligibility of Dutch English by having the listeners transcribe two sentences. Although this is one of the methods frequently used in this type of research, we feel the need for a more thoroughly tested approach to measuring intelligibility. Wang's (2007) transcription test using grammatically correct but semantically nonsensical sentences seems promising in that it really tests the intelligibility only, and leaves out certain factors that might aid intelligibility such as relying on the context of the message; yet this method needs further testing in other contexts and using more elaborate tasks and materials.

To test the comprehensibility and interpretability, we had the listeners summarise the meaning of the message they had heard. We chose this approach because it seems to us that just asking participants how well they understood the message, without having any means of knowing whether they really understood as much (or as little) as they claim will not produce the most reliable results. Yet, in order to measure comprehensibility and interpretability, a more thoroughly tested definition of both aspects and alternative methodology (e.g. multiple-choice questions) and more complex texts or a combination of textual types could produce alternative results. As far as the Dutch speakers are concerned, it remains to be seen whether a lower level of English and hence a heavier accent would lead to the same results. The moderate accent we used is only obtained by less than a quarter of the Dutch population (Sociaal en Cultureel Planbureau, 2005). Although 87% of the Dutch population claim to be able to hold a conversation in English (European Commission, 2006), the majority will have a lower level of English proficiency than the speakers we used (cf. van Onna & Jansen, 2006).

In the case of speakers with other mother tongues, such as French, Spanish or Chinese Mandarin, future studies of the effect of their Englishes might reveal interesting results. Coupland and Bishop (2007) show for example that French English evoked more positive attitudes than Spanish English, and Spanish English evoked more positive attitudes than German English. As far as the listeners in this study are concerned, we only investigated effects on British listeners, but it would be worthwhile to see if our results can be extrapolated to other groups of listeners: L1 speakers (e.g. Americans; Australians), L2 speakers (e.g. Indians) and FL speakers (e.g. Danes; Germans; Frenchmen). It might well be the case that for these listeners the results would be different, either because they are less familiar with different varieties of FL English, or because their first language may be similar or different to Dutch.

2.5 SUPPORTING INFORMATION

2.5.1 The text of the stimulus

Good afternoon, this is Jane Smith of SMG Incorporated speaking. We are one of the most well-known Dutch asset management businesses and we are looking to expand our market in Great Britain. For this reason we would like to discuss a possible partnership with your organization. Would you perhaps be interested in what SMG could mean for your organization? If this is the case, I can tell you more about our company, or we could make an appointment to discuss it in more detail.

Speech sample links: verbal guises
The speech sample links below redirect to one speech sample per accent (standard British English, slight Dutch-accented English, moderately Dutch-accented English).

- standard British English
  https://surfdrive.surf.nl/files/index.php/s/9sIarmdYJ5H6SzO
- slight Dutch-accented English
  https://surfdrive.surf.nl/files/index.php/s/oPePa5kjf5284Cz
- moderate Dutch-accented English
  https://surfdrive.surf.nl/files/index.php/s/YF4MXPklthbKEHv

2.5.2 Phrases used for the intelligibility, comprehensibility, and interpretability task

1. Good afternoon, this is Jane Smith of SMG Incorporated speaking.
2. We are one of the most well-known Dutch asset management businesses and we are looking to expand our market in Great Britain.

Intelligibility task: Please transcribe the sentences orthographically.
Comprehensibility task: Please paraphrase the message.
Interpretability task: Please indicate the purpose of the message.

3
REFINEMENT OF THE MATCHED-GUISE TECHNIQUE

ABSTRACT

Finding and selecting representative matched guises of native and non-native English accents is challenging, because people cannot be both a native and a non-native speaker of one language. In order to test a method to establish representative matched guises, matched-guise samples were recorded for the non-native accent, Dutch English, and standard British and American English accents. Four speakers, pre-selected by language experts, each produced all three accents. The samples were then evaluated on degree of 'nativeness' and 'standardness' by three 'linguistically naive' listener groups: 40 native British English speakers evaluated the British English samples, 40 native American English speakers evaluated the American English samples, and 40 native Dutch speakers evaluated the Dutch English samples. Despite the language experts' assumptions regarding the representativeness of the accents produced by the four matched-guise speakers, only two of the four speakers were deemed representative speakers of all three accents. This suggests that representative matched guises for the purposes of (non-)native English accents research can be achieved, but that speaker evaluations by 'linguistically naive' native listeners are necessary to ensure valid matched-guise selections.

This chapter is based on:
Nejjari, W., Gerritsen, M, Planken, B. and Van Hout, R. (2019).
Refinement of the matched-guise technique for the study of the effect of non-native accents compared to native accents. Lingua, 90-105.

3.1 INTRODUCTION

The purpose of our study was to establish a matched-guise selection method to study responses to native and non-native English accents. Accents are one of the first speech elements observed by listeners and can arouse strong positive or negative evaluative responses. Therefore, accents can have a significant impact on the attitudes towards, as well as the (perceived) comprehensibility of, a speaker (e.g. Nejjari et al., 2012; Hendriks, van Meurs & de Groot 2017; Orikasa, 2016; Rindal & Piercy, 2013; Tokumoto & Shibata, 2011).

The research reported here is a preliminary study conducted to establish matched guises for a larger speaker evaluation and speech comprehension[4] research project on non-native listeners' (Dutch, German, Spanish, Singaporean) responses to Dutch-accented English, standard British English and standard American English. There has been considerable debate on the status of English in Netherlands (e.g. Gerritsen et al., 2016; Edwards, 2016) and a number of experiments have been conducted to establish how native speakers of English view Dutch-accented English (e.g. Nejjari et al., 2012). However, the effects of Dutch-accented English on the groups of non-native speakers of English that Dutch people communicate with in English in professional contexts, which generally concern highly educated speakers who use English as a language and/or lingua franca in an international, professional context, e.g. in academia or business, have not often been researched (but see e.g. Hendriks, van Meurs & Reimer, 2018).

We aimed to assess how, compared to highly educated, native and standard speakers of English, highly educated Dutch speakers of English are perceived by other highly educated, non-native English speaker listeners in terms of the status, solidarity and dynamism (speaker evaluations) that their accent evokes as well as the extent to which their speech is understood (speech comprehension). In order to measure listeners' responses to only the three accents, and not to individual speakers, the matched-guise technique was applied to create speech samples for the larger research project. The present study describes the process used to establish representative matched guises that could be used in our research project on the response to native and non-native accents of English, and consequently also in

4 Speech comprehension in this chapter refers to what is defined as speech understandability in this thesis.

research involving other native and non-native accents and listener groups with varying education levels and backgrounds.

The two native and standard English accents (British and American) were selected because they are varieties that many learners of English around the world aspire to master and that are well documented in terms of their linguistic features (Nejjari et al., 2012; Gussenhoven & Broeders, 1997). Generally, a standard language is defined as a stable and established written and spoken language variety accepted by a speech community as the standard variety of one language (Grondelaers & Kristiansen, 2013; Milroy, 1992; Meyerhoff, 2011; Smakman & Barasa, 2017). This standard is commonly present in media, education, academia, and government (Smakman & Barasa, 2017). However, there is also debate on the concept of standard language or 'standardness', for example, with regard to the conceptualization of linguistic standardness in multilingual and diglossic societies (Smakman & Barasa, 2017). In addition, there is debate on how linguistic features of varieties show significant variance within European and Western standard languages (Coupland & Kristianson, 2011; Grondelaers & Kristiansen, 2013; Smakman, 2006, 2012). Even the concept of 'standard' in the sense that it represents a static, uniform representation of a language is debated, because some researchers believe the concept of 'standardness' represents a linguistic ideal or more extremely put a 'myth' that changes over time and thus does not reflect linguistic reality whatsoever (Lippi-Green, 1997:p. 44). Despite these points of discussion, it remains important to tap into listeners' concept of standardness when establishing matched guises since it offers researchers insights into listeners' current intuitions of what a standard accent sounds like and thus what 'standardness' is in a relevant speech community. Furthermore, this information can be used to assess the suitability of speakers in the type of speaker evaluation research that the present study is a part of (see Grondelaers & van Hout, 2015 for an overview of this discussion; Trudgill, 2001; Wolfram & Schilling-Estes, 2006).

Secondly, the non-native English accent Dutch English was selected because we were interested in studying the speech perceptions towards, and comprehension of, a non-native English accent from a country that generally has a good reputation in terms of its citizens' English fluency, being ranked first in EP's international ranking of English fluency by non-native speakers (EP English Proficiency Index, 2017). We wanted to investigate whether this ascribed good fluency would result in speech perceptions and levels of comprehension for Dutch English that are comparable to those for the two native English accents, which might imply in practice that Dutch English speakers no longer require accent training (for an overview of the discussion on Dutch English, see Gerritsen et al., 2016).

3.1.1 Speaker evaluation research and the matched-guise technique

Speaker evaluation research has long been aimed at understanding the impact languages and accents can have on the degree of understanding of, and attitudes towards, people speaking specific language varieties. A number of researchers have assessed how English is understood and perceived by speaker groups with different language backgrounds (Nejjari et al., 2012; Hendriks et al., 2017; Orikasa, 2016; Rindal et al., 2013; Tokumoto et al., 2011). Commonly used research methods aimed at studying the evaluations of speakers and the comprehension of languages and accents are the verbal-guise and the matched-guise techniques. Verbal guises are speech samples produced by native speakers who each produce their native language and no other (e.g. Nejjari et al., 2012; Hendriks et al., 2017, Hendriks, et al. 2018). A verbal-guise technique provides listeners with speech produced by several native speakers and is employed to collect speech samples that sound as natural and authentic as possible, and simultaneously allows researchers to control the content produced by the verbal-guise speakers. Similarly, the matched-guise technique is aimed at controlling the content listeners are exposed to, but differs in that one speaker produces all language varieties tested. Listeners are not aware of this fact and, unlike with verbal guises, this ensures they respond to the accent under study and not to the individual speaker. What is also of importance in the successful application of the matched-guise technique, and indeed the verbal-guise technique, is that listeners are able to correctly identify the language or accent they are required to respond to in order to guarantee that the gathered data reflects the associations listeners have with the target languages or accents (Preston, 1989).

The matched-guise technique is mostly applied to compare languages or accents of a language, and generally uses speakers who have learned more than one language (e.g. bilinguals or multilinguals) to such an extent that they are viewed as native speakers of these languages by native speakers (e.g. Baker, 1992; Cavallaro, Ng & Seilhamer, 2014; Dalton-Puffer Kaltenboeck & Smit, 1997; El-Dash & Busnardo, 2001; Garrett, 2010; Giles, 1970; He & Zhang, 2010; Lambert et al., 1960; Lindemann, 2003; Purnell, Idsardi & Baugh, 1999; Schüppert, Haug Hilton & Gooskens, 2015). Although the matched-guise technique has been used to study the response to many languages, dialects, and accents (e.g. Cavallaro, Ng & Seilhamer, 2014; El-Dash & Busnardo, 2001; Giles, 1970; Lambert et al., 1960; Purnell et al., 1999; Schüppert et al., 2015), and despite recent calls to apply the matched-guise technique in studies aimed at comparing non-native English accents with other non-native and native English accents, the use of the matched-guise technique remains rare in this type of speaker evaluation research (see Mai & Hoffmann, 2014).

The fact that the matched-guise technique is relatively rarely used in this context might be connected to what Garrett (2010) refers to as the 'mimicking-authenticity question' in his discussion of the benefits and limitations of the matched-guise technique. Garrett points out that when more than two (varieties of) languages are produced by one speaker, this can negatively reduce the accuracy of speech production. This is because it can lead to speech samples that sound somehow different, even though they are perceived as representative. Representativeness in the context of this study refers to the extent to which an accent is accepted as being authentic. This means that the accent produced is accepted by listeners as genuine, produced by a native speaker of that accent, and therefore is not performed, 'emulated' or 'put on' by a speaker who is not a native speaker of a language variety, a definition based on Guy and Cutler (2011: p. 139). The concept of representativeness provides an additional challenge when matched guises are required for speech evaluation research focused on combinations of native and non-native accent varieties, because a person who is a native speaker of, for example, British English cannot be a 'native' speaker of a non-native variety of his/her own language as well. This challenge to find representative matched guises, pointed out by amongst others Garrett, is confirmed by researchers who have attempted to find single speakers to produce two or more native and non-native guises. Approximately two decades ago, for example, Dalton-Puffer et al. (1997), reporting on a speaker evaluation experiment involving Received Pronunciation, General American, and two types of Austrian non-native accents, noted their frustration in not being able to determine representative matched guises for their experiment. They stated that '[i]t is practically impossible to find speakers who are equally convincing in several guises. This means of course that variables like voice quality can be controlled only minimally.' (p. 117).

However, the fact that it is difficult to find a matched-guise speaker should not deter researchers from attempting to use the matched-guise technique, because it offers important benefits too (see above). The abovementioned example and the general discussion regarding the benefits and limitations of the matched-guise technique (see Garrett, 2010) highlight that the selection of matched guises for speaker evaluation research on non-native and native English accents requires a specific strategy to avoid inauthentic and unusual sounding speech samples and instead allows for speech samples that represent the tested accents.

Generally, researchers have tended to select matched-guise speakers themselves (e.g. Lambert et al., 1960; MacFarlane & Stuart-Smith, 2012; Purnell et al., 1999; Schüppert et al., 2015). These researchers are linguists and/or experts in a specific language who believe they can judge whether speakers represent the targeted languages or

accents. In some cases, researchers have asked others to assess the quality of their guise selection. For example, in the study by Purnell et al. (1999), one of the researchers was the matched-guise speaker who produced the three dialects being investigated: African American Vernacular English (American English dialect associated with African Americans), Chicano English (American English dialect associated with Latino Americans), and Standard American English (American English dialect associated with European Americans). This speaker was familiar with all three dialects and was defined by other researchers in the team as a tridialectal speaker. The matched guises were mixed with speech samples produced by 20 random native speakers of one of the three dialects. Purnell et al. (1999) then investigated whether identification for the three guises was possible. The speech samples were randomly presented to 382 native speakers of English and 39 non-native speakers of English who had to identify the ethnicity of the speakers (multiple choice answer options were: African American, Hispanic American or European American). The results showed that the matched guises were accepted by at least 84% of all listeners for all three dialects, providing support that the initial selection of matched-guise speakers had resulted in representative guises.

Yet another methodology was used to assess the representativeness of matched guises under study by Schüppert et al. (2015), who tested whether their Danish and Swedish matched guises actually represented the required degree of nativeness by means of a so-called 'voice parade': a sequence of five voice recordings in either Danish or Swedish that was presented to listeners. The voice parade was created using one matched-guise speaker who produced the guises for Danish and Swedish, and four recordings of native speakers. In total, five recordings of Danish (guise or native) were presented to 30 native speakers of Danish and five recordings of Swedish (guise or native) were presented to 15 native speakers of Swedish. These native speakers of Danish and Swedish or 'native listeners' were asked to choose the speaker they believed sounded non-native, meaning a speaker that sounded foreign. If the guises were not regarded as coming from a non-native above chance level, the guise was accepted as a valid guise. Similar to Purnell et al. (1999), Schüppert et al. (2015) also required listeners to choose which language was represented by each speech sample. The results of this matched-guise assessment method show that the native listeners generally agreed with the researchers' perceptions of what is a representative matched guise.

Although the abovementioned studies show that it is indeed useful to tap into the perceptions of non-linguists or 'linguistically naïve', native listeners, research focused on non-native Englishes requires a specific research strategy, because there

are no set standards for non-native varieties of English in terms of grammar, vocabulary, and accent. Investigating such accents therefore requires extra care on the part of researchers to establish the matched guises that are to represent specific accents. Similar to the abovementioned studies, finding suitable matched guises should involve assessments by native listeners who are familiar with the tested accent varieties and therefore can judge whether matched guises are indeed representative. The present study demonstrates the use of this approach, and employs the response to potential matched guises of the different accents by naïve, native listeners, that is, non-linguists and non-language specialists who are native speakers of British English, American English, and Dutch.

These listeners responded to the matched guises of the accent from their native language. For example, native British English listeners responded to standard British English matched guises while native American English listeners responded to standard American English matched guises. This approach was taken to guarantee familiarity with these accents, since native speakers of languages can generally form clear judgments of speech as being native or not, and standard or not, even though the perception of what is native and standard can vary (Grondelaers and van Hout, 2015). For the non-native accent, which is Dutch English in the present experiment, rather than having listeners assess the standardness of a matched guise, they were asked to assess the degree to which a matched guise sounded typical for a native speaker of Dutch speaking English. This was done because asking listeners to indicate the standardness of a Dutch English accent can be confusing since Dutch English is not an official English variety and has no established standardized grammar, vocabulary or accent. Establishing the typicality of the Dutch English guises in the present study was therefore important in assessing whether they could suitably represent the specific (Dutch) English accent under study.

3.1.2 Dutch English

Before we elaborate on the selection of the matched-guise speakers in the present study, we define a typical Dutch English accent. There have been several studies aimed at establishing what constitutes a typical Dutch English accent. Such studies have, for example, described the phonological characteristics of native speakers of Dutch speaking English (Collins & Mees, 2003; Gussenhoven & Broeders, 1997; van den Doel, 2006; van der Haagen, 1998; van Hout et al., 2018). Although the English spoken by Dutch people can differ depending on the regional origin of a speaker, Dutch-accented English appears to possess many universal features. For example, a typical Dutch English accent contains features that native speakers of Dutch and

others familiar with Dutch and Dutch English will recognize as such. For example, Dutch lacks the English dental consonants [ð] as in *this*, *mother*, *breathe* and [θ] as in *think*, *Martha*, *breath*. They are often pronounced as stop consonants, [d] and [t] respectively, by Dutch speakers of English. Dutch also lacks voiced fricatives and plosives in the coda, causing the voiced obstruents of English to be pronounced as their voiceless counterparts (e.g. *live*, *badge*, *bad*, *bag* will be said with [f, tʃ, t, k]) (Gussenhoven & Broeders, 1997).

However, despite the documented characteristics of a Dutch English accent, there is, to our knowledge, no defined Dutch English accent standard. This aspect of Dutch English accent makes it challenging to apply the matched-guise technique because, to assess the representativeness of a Dutch English accent sample, researchers need to tap into speakers' generally held consensus – rather than an established standard – on what a Dutch English accent sounds like.

3.1.3 Research question

To increase the likelihood that a valid selection of matched guise speakers could be made to study the perception of native versus non-native accents in the larger study, native speakers of British English, American English and Dutch were asked to evaluate whether four matched-guise speakers selected by the researchers had actually produced a representative accent for standard British English, standard American English, and Dutch English. Representative accents of standard American English, standard British English, and Dutch English in the context of the present study refer to accents that are accepted as being a native and standard British English or American English accent, and as an accent that originates from a native speaker of Dutch with a typical Dutch accent in English. The central research question was:

> *Can representative matched-guises be established for speaker evaluation research aimed at researching responses to non-native and native English accents?*

3.2 METHOD

The method section describes the approach taken to assess whether four matched-guise speakers, chosen by experienced (socio)linguists and English language specialists, were able to produce representative accents of three varieties of English, namely standard British English, standard American English, and Dutch English according to three linguistically naïve, native listener groups.

3.2.1 Listeners and questions

Three groups of speakers of English listened to and evaluated the matched-guise speakers who produced Dutch-accented English, standard British English, and standard American English accents. This experiment is focused on the usability of the matched-guise technique in speaker evaluation research focused on non-native accents of English, in this case Dutch-accented English, compared to native accents. Therefore, two native accents were included to compare the listener groups' responses. In the context of this study British English and American English were selected, which are both accents that many people around the world aspire to master and are well documented in terms of their linguistic features, and are both also most commonly used in Dutch education, business, culture and academia (Nejjari et al., 2012; Gussehoven & Broeders, 1997; van der Haagen, 1998).

The experiment involved 40 listeners per accent, resulting in a total of 120 listeners, one group for each accent. Each listener group consisted of listeners with the same language and country background: group 1 consisted of native speakers of British English (n=40) mainly born and raised in the Manchester area (UK), group 2 of native speakers of American English (n=40) mainly born and raised in the Northwestern part of the United States, and group 3 of native speakers of Dutch (n=40) mainly born and raised in the Southeastern part of the Netherlands. We assume that the responses to the matched guises were unlikely to have been affected by the regional background of the listeners, since native speakers with different regional backgrounds generally tend to have similar perceptions of what the native and standard variety of their language is (Grondelaers and van Hout, 2015). The questionnaire for the British English native listeners was in British English. The questionnaire for the American English native listeners was in American English (e.g. 'check the box' and not 'tick the box'). The questionnaire was in Dutch for the Dutch native listeners. All native listeners (N=120) were university students who did not have a background in linguistics and who had no knowledge of the purpose of the experiment (Table 1).

Each listener was asked to evaluate 24 speech samples (see: *3.2.2 Speakers and speech samples*) by answering two questions about each sample. These two questions were aimed at assessing (1) the extent to which the listeners believed a speaker was a native speaker of a target language variety, the 'nativeness' question, and (2) to which extent they believed the speaker had a representative accent of the target variety: the 'standardness' (or typicalness) question. The two questions served as a way to make a distinction between what is considered native (British English; American

Table 1. The three listener groups.

Location of the listeners	Number of listeners	Mean age (SD)	Males	Females	Education level obtained
Native Speakers of British English: The University of Manchester, UK	40	23 (4.00)	33%	67%	Undergraduate/ Bachelor
Native Speakers of American English: Montana State University, US	40	25 (6.45) 3 listeners were above 40 years of age	40%	60%	Undergraduate/ Bachelor
Native Speakers of Dutch: Radboud University Nijmegen, The Netherlands	40	23(2.69)	43%	57%	Undergraduate/ Bachelor

English; Dutch) and what is considered a standard accent (standard British English; standard American English) or a typical language variety (Dutch English). A speaker may be identified as a native speaker of British English, but his or her social or regional accent may be viewed as being too strong or deviant to be evaluated as a typical speaker within the limits of the target accent. Social and regional variation is relevant of course, but our aim in this investigation was to end up with speakers representing the target varieties beyond discussion.

For each speech sample, The listeners had to indicate on 7-point Likert scales (1=completely disagree; 7=completely agree; 4=neither agree nor disagree) to what extent they agreed with the two statements, for example, for the British listener group, these were:

1. The speaker is a native speaker of British English.
2. The speaker is a standard speaker of British English.

For the American listeners, the first and second statement referred to American English instead of British English. For the Dutch listeners, the questionnaire was in Dutch and the first statement referred to Dutch ('De spreker is een moedertaalsperker van het Nederlands' in English: 'The speaker is a native speaker of Dutch') and the second statement was: The speaker has a typical Dutch English accent (in Dutch: 'De spreker heeft een typisch Nederlandse uitspraak in het Engels'). A speaker needed to achieve a mean score of at least 4 on the 7-point scales measuring 'nativeness' and 'standardness' (1=completely disagree; 7=completely agree; 4=neither agree nor disagree) to be considered a representative matched-guise speaker. A mean

score of four was selected as a threshold because it signifies a neutral, not negative or positive, attitude toward a speaker in terms of 'nativeness' and 'standardness' (or typicalness). Also, to the best of our knowledge, there is no validated threshold for nativeness and standardness. As explained earlier, the listeners were assumed to have clear perceptions of what sounds native and standard (or typical), even though assessments might vary slightly between individuals (Grondelaers and van Hout, 2015). This has also been demonstrated by Smakman (2006) who studied the phonetic features of speakers of standard Dutch. He found both consistent and deviating features in standard Dutch speech, which had been selected by listeners who were native speakers of Dutch (2006). This could mean that rather than an absolute phenomenon, 'nativeness' and 'standardness' are continuous phenomena that are best measured as a range. Therefore, the range for representativeness should be set at values ranging from 4 (neutral) to 7 to allow for variation within a given range to emerge from the data. A threshold set at any value above 4, for example 5.5, would imply that any scores between 4 and 5.5 are not native and standard enough. However, there is no valid reason to set the threshold at that exact point, because it cannot be proven that 5 would not be high enough a score to achieve representative speech samples, since it is on the positive side of the 7-point scale, even if a 5 score could be significantly lower statistically speaking than 5.5. Therefore, the threshold should start at a non-negative point which is neutral, and in our case was a score of 4.

The listeners were also asked to provide bio-data (e.g. gender, age, education, and e-mail address). The experiment was conducted individually or in groups in lectures, but always in the presence of the first author or a research assistant. For the British listeners, the experiment was conducted by a research assistant at Manchester University in 2013/2014. For the American listeners, the experiment was conducted by a research assistant at Montana State University in 2014. The data from the Dutch listeners were collected at Radboud University Nijmegen in 2013 by the first author and a research assistant.

3.2.2 Speakers and speech samples

Each listener group evaluated 24 speech samples but eight of these were the matched-guise speech samples (see *section 3.5*). In addition, they were presented with 12 filler and four control speech samples. These were added to prevent listeners from deducing they were listening to matched guises for a specific accent. Furthermore, control samples were added to allow the researchers to compare the matched guises to natural native speakers of the three accents.

3.2.3 The matched-guise speech samples

Four matched-guise speakers were selected by the researchers to produce the three accents (Table 2). The four matched-guise speakers were all linguists and English language specialists with specific linguistic expertise in terms of Dutch/British/American English (accents). In addition, one of the speakers was regarded as bilingual in Dutch and British English, because she has acquired Dutch and British English to such an extent that she is generally informally as a native speaker of these languages by native speakers. The other three speakers were native speakers of Dutch with reputations for being native-like speakers of native, standard British or American English by native speakers of British and American English. Therefore these three speakers can also be viewed as bilingual. Four matched-guise speakers were selected to better guarantee that reactions to the speech samples based on accent and not on personal characteristics such as voice quality and speech rate, which can result in a more positive or negative evaluation of an individual speaker on the basis of perceptions of these aspects and not of accent. Having more than one matched-guise speaker produce the three accents thus allowed for a more reliable determination of the representativeness of the guises.

Each matched-guise speaker was asked to read out a text for each of the three accents: (1) standard British English, (2) standard American English, and (3) the English accent of highly educated native speakers of Dutch. The Dutch English accent of a highly educated speaker was selected, because these speakers represents the segment of Dutch society most likely to use English as a language and/or lingua franca in an international, professional context, e.g. in academia or business. To produce the samples, the matched-guise speakers read out the two texts at least

Table 2. Four matched-guise speakers with two speech samples per accent; MG = matched-guise speaker; 1,2,3,4 = speaker number; a,b = sample a or b.

Speakers	British English samples	American English samples	Dutch English samples
Four Matched-guise speakers (MG)	MG1a MG1b	MG1a MG1b	MG1a MG1b
	MG2a MG2b	MG2a MG2b	MG2a MG2b
	MG3a MG3b	MG3a MG3b	MG3a MG3b
	MG4a MG4b	MG4a MG4b	MG4a MG4b

two times. Subsequently, samples of 10 seconds each were selected as the final matched-guise speech samples. The researchers selected the speech samples that sounded the most natural and had good sound quality. This was done to see if the responses to the matched-guise speakers consistently achieved a mean score of at least 4 to further ensure the representativeness of the matched-guises. This resulted in six speech samples produced by each matched-guise speaker (see Table 2).

3.2.4 The control and filler speech samples

In addition to matched-guise speech samples, each listener was asked to evaluate four control speech samples. These were speech samples in the listener groups' native language produced by natural native speakers of that particular accent (Controls). In addition, filler speech samples were used to prevent listeners from deducing our research aim and to make the task more attractive by offering another language variety, each listener group also had to evaluate eight matched-guise speech samples of one of the other accents, and four native speech samples of that same accent produced by native speakers of that particular accent (Matched-guise Fillers and Native Fillers). This meant that, for example, the British listeners, who were selected to evaluate the four matched-guise speakers' standard British English, were presented with the eight British English matched guises as well as the eight

Table 3. Speech samples per listener group: each listener evaluated 24 speech samples, 12 from his/her own native language (8 Matched guises and 4 Controls) and 12 from another language variety (8 Matched-guise Fillers and 4 Native Fillers).

Listener group:/ Speech samples:	British English listeners	American English listeners	Dutch English listeners
Own variety	British English	American English	Dutch English
Matched guises	8 Matched guises	8 Matched guises	8 Matched guises
Controls	4 Controls	4 Controls	4 Controls
Other variety	Dutch English	Dutch English	American English
Matched-guise Fillers	8 Matched-guise Fillers	8 Matched-guise Fillers	8 Matched-guise Fillers
Native Fillers	4 Native Fillers	4 Native Fillers	4 Native Fillers

Dutch English matched guises. Also, they had to evaluate four native speakers of British English speaking British English and four native speakers of Dutch speaking Dutch English. Table 3 illustrates the control and filler sample structure per listener group.

3.2.5 Selection of the control speakers

The native speakers of British English, American English, and Dutch that were selected to act as controls to the matched guises were selected on the basis of their gender and professional background. Two males and two females were selected per accent. The native speakers of British English and American English all had teaching experience in English and were trainers and/or editors of English. In addition, they were all informally characterized as native and standard speakers. The native speakers of Dutch were all language and communication professors and were used to publishing and communicating in English in an international, professional setting, which is the variety of Dutch English accent assessed in this study, as explained earlier.

Three speakers who produced British English were born and raised in England in the London area, and the fourth native speaker was originally from Manchester. Three of the four speakers were English teachers at university language and communication centers, had obtained at least a bachelor's degree in the UK, and were between the ages of 50 and 60. The fourth speaker was a master student as well as a teacher's assistant and was 25 years of age.

Two of the four native speakers of American English were born and raised in the US in New-Hampshire, obtained their degrees (at least bachelor's degree) in the US and worked as English trainers at university language and communication centers. The third speaker was born and raised in different cities on the East coast of the US and worked as a language lecturer. The fourth speaker was born and raised in the Netherlands by a Dutch mother and American father and was raised bilingually. He obtained his master's degree in the Netherlands and worked as a trainer, translator and editor of English at a university language and communication center. All American English speakers were between the ages of 34 and 65.

The native speakers of Dutch who produced Dutch English speech samples were all associate or full professors in the areas of language and communication at Radboud University. They were all between the ages of 45 and 65, born and raised in the southeast and south of the Netherlands, and all had taken and passed a Cambridge University English exam at C2 level, which according to the European Framework of Reference is the highest English language fluency level a learner can acquire (Council of Europe, 2017).

3.2.6 The speech sample texts

The 24 speech samples were based on two texts that were each approximately 1,000 words about the development of technology and robotics. The two texts were taken from the Cambridge IELTS academic English reading exam (British Council, 2017) and an online article (Menzel, 2017) both discussing animals, such as geckos, that inspire technological developments. The Cambridge IELTS academic English reading exam was used as a text resource because this internationally acknowledged exam is designed to use general interest topics to assess the English language skills of highly educated examinees who aim to test whether their language skills are sufficient for an academic environment, such as a university degree programme or an academic position. The online article was found on the basis of a reference in the reading exam to the MIT research project involving animals and technology and used to write the speech sample texts that were of a high quality and focused on a general interest topic.

The speech samples discussed the same topic, but the text of the speech samples differed in most cases to avoid exposing listeners to exactly the same segment 24 times, which would have increased the chances of diminished task engagement, and as a result might have influenced the listeners' responses in the sense that listeners might have been tempted to not concentrate as well as they could have. The speech samples were randomly mixed in two sequences, creating two versions of the questionnaire for each listener group.

3.2.7 Design

The independent variables were the matched-guise accents by our four speakers: British English, American English, and Dutch English plus the additional four control samples per variety from native speakers. The dependent variables were the evaluations by the native speakers of each speakers' degree of 'nativeness' and 'standardness'.

3.2.8 Procedure

Listeners were asked to evaluate 24 speech samples by different speakers in a lecture room where they were taking a course. The precise details of the purpose of the experiment were not shared with the listeners, but the questions of the questionnaire were explained to the listeners in terms of what was required of them. It was explained that we were very curious about the way the speakers came across. Listeners needed approximately 20 minutes on average to complete the questionnaire. The speech samples were played on the lecture room computer's central sound

system. Listeners were allowed to listen to each speech sample more than once and were asked after each speech sample was played whether this was necessary. Five Dutch listeners asked to listen to the first two of the 24 speech samples twice, which were played two times in the central lecture room, because they were distracted by something or the volume was set too low for them.

3.2.9 Statistical treatment

SPSS 23 was used to calculate means, standard deviations, standard errors, and interrater reliability (Intraclass Correlation Coefficient to measure the consistency between measures of nativeness and standardness). Confidence intervals were also included at a 95% level to provide information on how likely our sample contained the population's actual means.

3.3 RESULTS

3.3.1 Reliability

Intraclass correlation coefficients (ICCs, with speakers and listeners as random factors) to measure interrater reliability were calculated for each respondent or listener group (n=40). A .90 ICC was considered a sufficient score to accept the validity of the total listener groups' evaluations. All ICCs were above .90. For the British English listener group a .90 ICC was achieved at 15 listener evaluations for the question whether the speaker was a native speaker of British English, and an ICC of .90 achieved at 19 listener evaluations for the question whether the speaker had a standard British English accent. For the American English listener group a .90 ICC was achieved at 19 listener evaluations for the question whether the speaker was a native speaker of American English, and an ICC of .90 achieved at 22 listener evaluations for the question whether the speaker had a standard American English accent. For the Dutch listener group a .90 ICC was achieved at 16 listener evaluations for the question whether the speaker was a native speaker of Dutch, and an ICC of .90 achieved at 12 listener evaluations for the question whether the speaker had typical ('standardness') Dutch English accent. These results indicate that the number of listeners in this experiment was more than sufficient to generate reliable evaluations.

3.3.2 British English matched guises

Figure 1 and Table 4 display the mean evaluations for the eight speech samples by four British English matched-guise speakers (MG:1a,1b,2a,2b,3a,3b,4a,4b) and the

Figure 1. Means (Min=1; Max=7) and confidence intervals (+/- 2 standard error) for 'nativeness' of British English (NativenessBritishEng) and 'standardness' of British English (StandardnessBritishEng) speech samples (matched guises: MG1a,b;2a,b;3a,b;4a,b; controls: NS1,2,3,4) rated by native listeners.

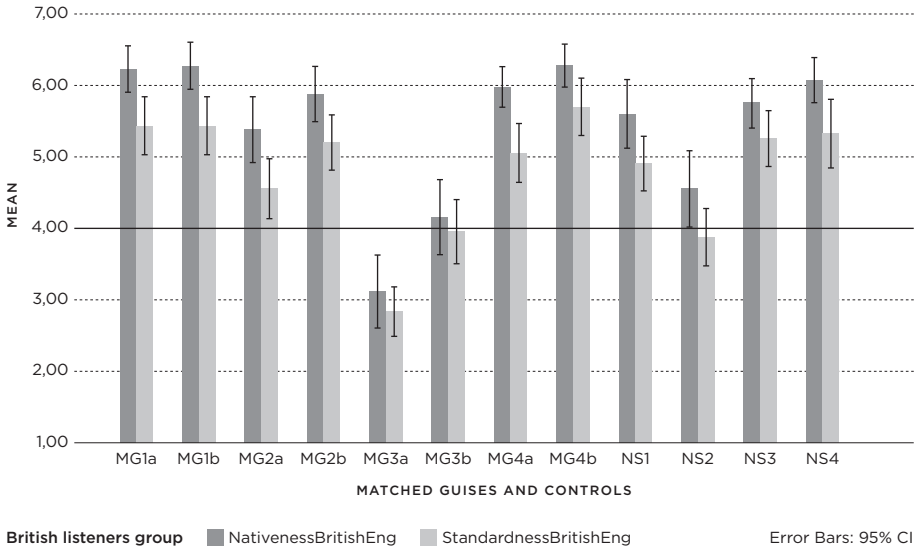


Table 4. Means (Min=1; Max=7) and standard deviations (SD) for British English matched guises (MG:1a,1b,2a,2b,3a,3b,4a,4b) and British English controls (NS:1,2,3,4) rated by native listeners.

Question/ Speaker	The speaker is a native speaker of British English	The speaker has a standard British English accent
	Mean (SD)	Mean (SD)
MG1a	6.23 (1.03)	5.43 (1.28)
MG1b	6.28 (1.04)	5.43 (1.28)
MG2a	5.38 (1.44)	4.55 (1.34)
MG2b	5.88 (1.20)	5.20 (1.20)
MG3a	3.10 (1.58)	2.83 (1.08)
MG3b	4.15 (1.66)	3.95 (1.41)
MG4a	5.98 (.89)	5.05 (1.30)
MG4b	6.28 (.93)	5.70 (1.26)
NS1	5.60 (1.50)	4.90 (1.19)
NS2	4.55 (1.66)	3.88 (1.26)
NS3	5.75 (1.10)	5.25 (1.21)
NS4	6.08 (1.00)	5.33 (1.51)

four speech samples by the four native speakers of British English or the British English controls (NS:1,2,3,4). The results show that three out of four matched-guise speakers were able to produce representative native and standard British English accents (MG:1a,1b,2a,2b,4a,4b). Their average scores were above 4 for both samples, which was the set minimum mean score required for all speech samples to be accepted as native and/or standard enough (See: *3.2.1 Listeners and questions*). Moreover, their means fit the range found for three of the British English controls ('nativeness' 5.60-6.08; 'standardness' 4.90-5.33). Matched-guise speaker three failed to reach a mean of 4 for 'nativeness' and 'standardness' for *sample a* and for 'standardness' of *sample b*, although the mean of 4 is included in its 95% confidence interval. Since matched-guise speaker three (MG:3a,3b) was clearly not able to meet the required degree for representativeness in two out of four means, she was not accepted as a representative guise. Finally, three British English controls (NS:1,3,4) were perceived by native listeners as being native speakers of British English with a standard British English accent. British English control 2 (NS2) had a mean of 3.88 for 'standardness', thus sounding native enough, but not unequivocally standard.

3.3.3 American English matched guises

Figure 2 and Table 5 display the mean evaluations for the eight speech samples by the four American English matched-guise speakers (MG:1a,1b,2a,2b,3a,3b,4a,4b) and the four speech samples by the four native speakers of American English or the American English controls (NS:1,2,3,4). The results show that two matched-guise speakers were able to produce representative native and standard American English accents (MG:3a,3b,4a,4b). Moreover, their scores fit the range found for the American English controls ('nativeness' 4.98-5.85; 'standardness' 4.63-5.83). Matched-guise speaker one (MG:1a,1b) achieved means below 3.5 for 'nativeness' and 'standardness' for both samples. Matched-guise speaker two (MG:2a,2b) was able to produce representative guises for both 'nativeness' and 'standardness' in *sample a* and for 'standardness' in *sample b*. Although, she achieved a mean of 3.83 for degree of 'nativeness' in sample b, the confidence interval of the standard error includes the scale mean of 4, thus making matched-guise speaker 2 a representative guise after all. All four of the American English controls (NS:1,2,3,4) achieved means above 4, and were thus perceived by the native listeners as representative native speakers of American English with a standard American English accent.

Figure 2. Means (Min=1; Max=7) and confidence intervals (+/- 2 standard error) for 'nativeness' of American English (NativenessAmericanEng) and 'standardness' of American English (StandardnessAmericanEng) speech samples (matched guises: MG1a,b;2a,b;3a,b;4a,b; controls: NS1,2,3,4) rated by native listeners.

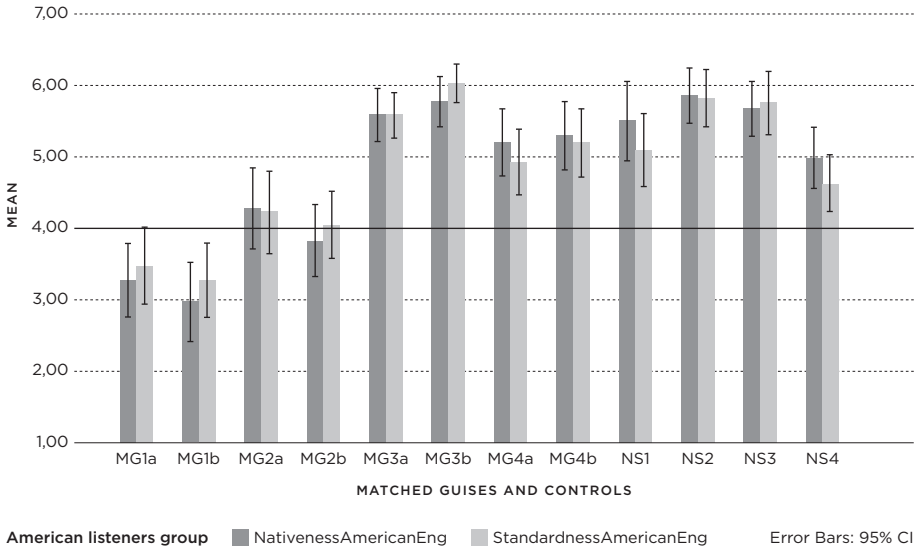


Table 5. Means (Min=1; Max=7) and standard deviations (SD) for American English matched guises (MG1a,b;2a,b;3a,b;4a,b) and American English controls (NS1,2,3,4) rated by native listeners.

Question/ Speaker	The speaker is a native speaker of American English	The speaker has a standard American English accent
	Mean (SD)	Mean (SD)
MG1a	3.28 (1.62)	3.48 (1.69)
MG1b	2.98 (1.75)	3.28 (1.62)
MG2a	4.28 (1.78)	4.23 (1.79)
MG2b	3.83 (1.55)	4.05 (1.48)
MG3a	5.58 (1.17)	5.58 (.98)
MG3b	5.78 (1.12)	6.03 (.83)
MG4a	5.20 (1.45)	4.93 (1.40)
MG4b	5.30 (1.51)	5.20 (1.49)
NS1	5.68 (1.19)	5.75 (1.25)
NS2	4.98 (1.33)	4.63 (1.37)
NS3	5.85 (1.23)	5.83 (1.26)
NS4	5.50 (1.72)	5.10 (1.60)

Figure 3. Means (Min=1; Max=7) and confidence intervals (+/- 2 standard error) for 'nativeness' of Dutch (NativenessDutchEng) and typical Dutch English (Standardness-DutchEng) speech samples (matched guises: MG1a,b;2a,b;3a,b;4a,b; controls: NS1,2,3,4) rated by native listeners.

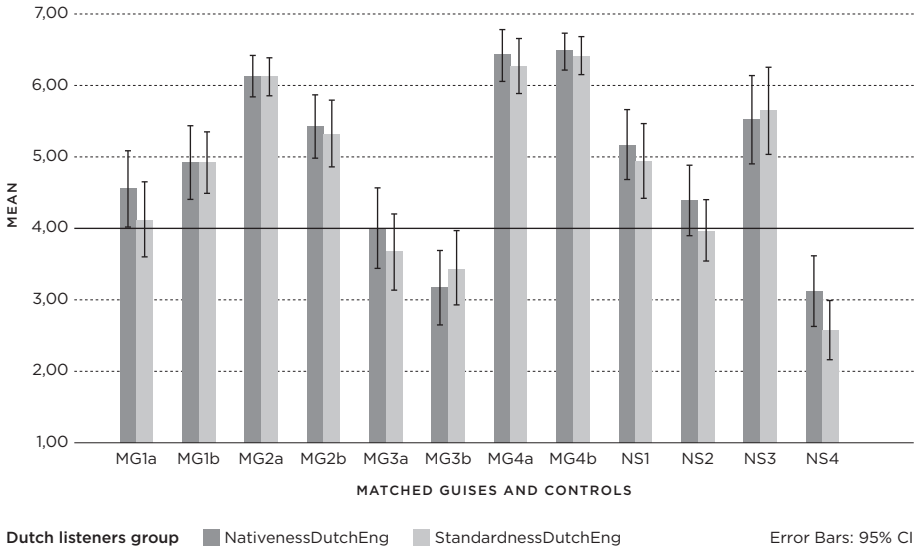


Table 6. Means (Min=1; Max=7) and standard deviations (SD) for Dutch English matched guises (MG1a,b;2a,b;3a,b;4a,b) and Dutch English controls (NS1,2,3,4) rated by native listeners.

Question/ Speaker	The speaker is a native speaker of Dutch	The speaker has a typical Dutch English accent
	Mean (SD)	Mean (SD)
MG1a	4.55 (1.68)	4.13 (1.64)
MG1b	4.93 (1.61)	4.93 (1.35)
MG2a	6.13 (.91)	6.13 (.82)
MG2b	5.43 (1.38)	5.33 (1.44)
MG3a	4.00 (1.77)	3.68 (1.65)
MG3b	3.18 (1.63)	3.45 (1.61)
MG4a	6.43 (1.13)	6.28 (1.15)
MG4b	6.48 (.78)	6.43 (.81)
NS1	5.18 (1.53)	4.95 (1.65)
NS2	4.40 (1.55)	3.98 (1.33)
NS3	5.53 (1.93)	5.65 (1.96)
NS4	3.13 (1.54)	2.58 (1.30)

3.3.4 Dutch English matched guises

Figure 3 and Table 6 display the mean evaluations for the eight speech samples by four Dutch English matched-guise speakers (MG:1a,b;2a,b;3a,b;4a,b) and the four speech samples by the four native speakers of Dutch speaking English or the Dutch English controls (NS:1,2,3,4). The results show that three matched-guise speakers were able to produce representative Dutch English guises (MG:1a,1b, 2a,2b,4a,4b). Moreover, their mean scores fit the range found for two of the controls ('nativeness' 5.18-5.53, 'standardness' 4.95-5.65). Matched-guise speaker three (MG:3a,3b) failed to reach the required mean score of 4 for 'standardness' in *sample a* and for 'nativeness' and 'standardness' in *sample b*. In two cases the confidence interval did not include the required minimum mean score of 4 and was therefore not large enough (see Figure 3). Since matched-guise speaker three was not able to meet the required degree for representativeness in three out of four mean scores, she was not accepted as a representative guise. Finally, two Dutch English controls (NS:1,3) were clearly able to produce speech samples that were representative in terms of 'nativeness' and 'standardness'. With a mean score of 3.98 for typical ('standardness') Dutch English accent, Dutch English control 2 (NS2) failed to reach the required mean of 4. Also, Dutch English control 4 (NS4) achieved means below 3.15 for both degree of 'nativeness' and for degree of 'standardness'. Despite the language specialists' agreement that this speaker would be able to produce a representative Dutch English accent, the native listeners disagreed.

In sum, the results for all three accents show that the three native listener groups had a very clear perception of the representativeness of the speech samples, and thus the selected speakers. All four matched-guise speakers were perceived as representative speakers for at least one of the three accents. However, only two matched-guise speakers were accepted as representative speakers for all three accents (see Table 7).

Table 7. Summary representativeness matched guises for 'nativeness' and 'standardness'.

Matched-guise speaker	1	2	3	4
Representative native, standard British English accent?	Yes	Yes	No	Yes
Representative native, standard American English accent?	No	Yes	Yes	Yes
Representative typical Dutch English accent?	Yes	Yes	No	Yes

3.4 CONCLUSION AND DISCUSSION

The main aim of the present study was to determine whether representative matched guises can be established for speaker evaluation research focused on researching responses to non-native accents. It can be concluded that this is indeed possible.

3.4.1 Representativeness of matched-guise speakers

The results show that matched-guise speakers two and four proved native enough and standard enough to be able to representatively produce a standard British English accent and a standard American English accent, respectively, that is according to native speakers of these varieties. In terms of Dutch English, these two matched-guise speakers were also able to representatively produce an accent that the native Dutch listener group viewed as produced by a native speaker of Dutch with a typical Dutch English accent.

3.4.2 Representativeness of native speakers

The main purpose of this experiment was to test whether the selected matched-guise speakers were able to produce the three accents. However, by also including 'controls' or the speech samples of native speakers of British English, American English, and native speakers of Dutch speaking English, we also gained insight into how listeners perceived the representativeness or degree of 'nativeness' and 'standardness' of native speech.

What becomes apparent from the responses to the tested accents is that in general the listeners agree that the controls are representative of a native and standard accent for standard British English and standard American English, and representative of native speakers of Dutch speaking with a typical Dutch English accent. This indicates that there is consensus about what is representative, that is, what is native and standard in terms of accent. In other words, degree of 'nativeness' and 'standardness' appears to be measurable, because a clear pattern emerges from our listeners' responses: they express what is representative and what is not.

With respect to the native speakers of Dutch (i.e. the controls) a distinction is made by the listeners between speakers in terms of their representativeness. All controls are native speakers of Dutch, and therefore, would be expected to be considered representative in terms of 'nativeness'. However, one speaker (NS4) was not seen by the Dutch native listener group as being a native speaker of Dutch. Also, she was not regarded as being a speaker with a typical Dutch English accent, that is, as representative in terms of 'standardness' either. Her accent apparently did not contain

sufficient Dutch English accent features or cues that the listeners in the present study associated with a native speaker of Dutch speaking English with a typical Dutch accent. The question this evokes is what linguistics elements distinguished her accent from the other native speakers of Dutch who were deemed representative, and also what language she represented according to Dutch listeners. However, answering this question was beyond the scope of the present study.

The listeners' consistent evaluation of what is native and what is not is further supported by the results for the second matched-guise speaker (MG:2a,2b). She achieved mean scores for degree of 'nativeness' and 'standardness' of accent similar to that of the native speakers of British English. The second matched-guise speaker is actually bilingual: a native speaker of Dutch and British English who spent her teens in Britain. The British English listeners' assessment of this matched-guise speaker's representativeness matches their assessment of the British English controls, which lends further support to the notion that people have a clear and similar idea of when someone sounds native and standard or when someone does not, which our method was able to illustrate.

3.4.3 Matched-guise technique revisited for future research

Our method of matched-guise selection differed on several points from the traditional selection method. We incorporated the concept of measuring matched guises on degree of perceived 'nativeness' and 'standardness'. We measured native speakers' perceptions of the tested accents and provided information on the backgrounds of these native listeners, in order to ensure evaluations by linguistically naïve, native speakers, and create comparable listener groups. Secondly, it provided a blueprint of the criteria, content, and statistical procedures that were used to determine whether the guises selected were in fact representative. Furthermore, we included a more than sufficient number of listeners in the present study to ensure the results were reliable.

In addition, the criteria that were used to accept or reject a guise were also defined. We set a minimum mean score that had to be achieved for both degree of 'nativeness' and 'standardness'. We believe that by using this threshold, we were able to check the listeners' perception of the representativeness of speech ('nativeness' and 'standardness') and, based on this, select which speech samples were representative and which speech samples were not. This standard is a first attempt at trying to define how representativeness of guises can be measured, what acceptable speaker evaluation scores are, and provides useful information on how the selection of reliable guises can be conducted in future studies in a similar vein. In the case of

Dutch English, we have attempted to establish valid matched guises and have shown that this is possible.

Our experiment has some limitations. Firstly, only one non-native English accent, Dutch English, was assessed and our research question aimed to investigate assessments of only three accents. Accents from other parts of the world might have elicited different responses from 'naïve' native listeners. That is, listeners from societies that have lower exposure to native Englishes from the U.K. and the U.S. in their education, professional environment, or culture in general than listeners from the Netherlands, or who have higher exposure to other native and/or non-native varieties of English, might have responded differently to matched guises that are meant to represent their accent in English.

Secondly, perhaps different degrees of (non-native) accentedness should also be included in future research to see how to best establish representative matched guises for different degrees of (non-native) accentedness in speaker evaluation research.

Thirdly, any phonetic differences between the accents assessed in the present study were not individually established and analyzed. Instead, the researchers selected the speech samples (see section 3.5) that sounded representative and were comparable in terms of voice quality and speech rate for each accent and subsequently had the naïve, native listeners assess the accents as well. Even though the listeners were shown to have a very clear idea of what sounds native and standard, which concurs with the findings from other studies (see Grondelaers and van Hout, 2015 for an overview; Smakman, 2006), there is no information on which phonetic elements might have caused the listeners' assessment of a speech sample. The concept of representativeness in this study therefore has been established in general terms, that is, as the consensus amongst listeners that a speech sample was representative.

Fourthly, listeners were explicitly alerted to the variety tested in order to arouse their perceptions of what that specific accent should sound like, because they were literally asked to indicate to what extent they believed a speaker was a representative speaker of the tested variety (i.e. British, American or Dutch English). This might have influenced the listeners' evaluations of the speech samples since listeners were not required to indicate where they believed a speaker was from, and what the speakers' native language was, which would have further validated listeners' familiarity with the tested accent and its representativeness, if they were able to indicate the speaker's nationality and native language (variety).

Fifthly, the speakers read out a text and were not recorded in a natural conversation setting, which might have created more natural speech samples or stimuli.

We opted to have speakers read out a text, however, to ensure that similar content was used and listeners would only respond to the speakers' accent and not the content of the text.

Finally, the stimuli discussed the same topic and were part of the same story concerning animals and technology, but varied in terms of the exact text, which could have impacted the responses to the speech samples. We do realize that our choice to allow for variation in the speech samples' text content offers other methodological challenges. If, for example, more complex vocabulary was used in one speech sample compared to the others, this could create the impression of a more cultured and therefore perhaps a more likely standard speaker of a language. Even though the differences in terms of the stimuli texts could have impacted the responses to the stimuli, this was done to ensure optimal task engagement since 24 stimuli needed to be evaluated by the listeners.

To conclude, we have shown that linguistically naïve, native listeners offer essential information that should be included in speaker evaluation research that employs matched guises to assess speech perceptions of non-native and native language varieties. Our results indicate that these listeners did not always agree with the linguists' selection of matched-guise speakers. Since the native listeners who typically evaluate speech in speaker evaluation research tend not to be linguists, we suggest that researchers tap into the perceptions of linguistically naïve native listeners as part of their matched-guise speaker selection, so as to be able to create truly representative speech stimuli. Essentially, the experiment reported here provides an example of evidence-based matched-guise selection, and, as such, a replicable methodology for determining representative matched-guise speakers in future accentedness studies and speaker evaluation research. As we focused on Dutch English, and British and American English accents, further research is required to see whether matched guises can be established for other native and non-native English accent combinations and to further explore what 'representativeness' means and how it can be best researched. This is of importance, because teachers and learners of all (non-)native English varieties need to understand how they are perceived and to gain awareness of the effects their accents and speech might have in their interactions with others.

3.5 SUPPORTING INFORMATION

Speech sample links: matched guises and controls

The speech sample links below redirect to one speech sample per accent (standard British English, standard American English, Dutch English),and per matched-guise speaker (matched-guise speakers 1,2,3,4). For the controls, three speech sample links are available (for each accent one). The complete playlist of the speech samples is available at:

or:
https://www.youtube.com/playlist?list=PLXirys_Q318qXMakdaHNLRgN46cEifbK8

The remainder of the recorded speech samples are available on request.

4
DUTCH LISTENERS AND DUTCH-ACCENTED ENGLISH

ABSTRACT

In a matched-guise experiment, Dutch listeners responded to a Dutch English accent compared to a standard British and American accent in terms of speech understandability *(intelligibility, comprehensibility, interpretability*) and speaker evaluations *(status, affect, dynamism)*. Dutch listeners (N=392) evaluated these accents in three communication contexts: Lecture, Audio Tour, Job Pitch. Only context affected speech understandability: *comprehensibility* and *interpretability* were higher for the Lecture compared to the Audio Tour and the Job Pitch. Accent did not affect listeners' speaker evaluations for *affect* and *dynamism*, but evoked a lower *status* for Dutch-accented English than for the standard varieties. Context did not affect *status* but evoked more *affect* in the Audio Tour and in the Lecture than in the Job Pitch. Our main conclusion is that Dutch-accented English does negatively impact *status*, but not understanding, *affect* and *dynamism*. Moreover, context impacts understanding and *affect*.

This chapter is based on:
Nejjari, W., Gerritsen, M., Planken, B. & van Hout, R.
(Submitted and revision in progress). Dutch listeners' understanding and evaluations of Dutch, British and American English accents in three communication contexts.

4.1 INTRODUCTION

In recent years many have noted and debated the status of English in the Netherlands (e.g. Gerritsen et al., 2016; Edwards, 2016). This is due to the increase of professional contexts in the Netherlands that require Dutch people to communicate in English. For example, in English streams of university bachelor and master degree programmes and in specific multinational corporation settings (Gerritsen et al., 2016; BON, 2017; Bouma, 2016, 2018; Edwards, 2016; Huygen, 2017; Lizzini, Martijn, Munk, & De Regt, 2017; Nickerson, 2005; van Gaal, 2018; van Heest, 2018). Generally, the Dutch are considered to possess high proficiency in English, as non-native (L2) speakers (cf. the high Dutch ranking on the English Proficiency Index, EF (2018)). Despite the widespread use of English, and the measured high English proficiency, discussions in the Netherlands on the desirability of the use of English spoken by native (L1) speakers of Dutch, for example in a higher education context, often centres around the perceived English fluency levels of lecturers and their accents in English (e.g. BON, 2017; Bouma, 2016; Bronkhorst, 2015; van Gaal, 2018). Dutch students have indicated that the Dutch-accented English used by their professors has hindered their understanding of content, and resulted in a less effective lecture (e.g. Huygen, 2017). This perception was partly confirmed in an experiment on self-reported speech understanding of degrees of Dutch instructors' accents in English (Hendriks, van Meurs, & Hogervorst, 2016). However, actual speech understanding of Dutch-accented English by Dutch listeners should be assessed as well to study whether Dutch-accented English truly hinders the understanding of speech, which is what this study will investigate. Furthermore, the speaker evaluations Dutch-accented English evokes, also in relation with actual speech understanding, in a variety of communication contexts should be researched. This is needed to establish whether the professed effects of Dutch-accented English on Dutch listeners are correct, and if so, whether they apply to specific contexts. The results of this study can contribute to the general discussion held in the Netherlands on the desirability of English in many spheres of Dutch people's lives.

Accents are one of the first signals that are picked up to assess a speaker and the information conveyed. An accent indicates the national, regional, and socio-economic origins of a speaker and often evokes strong responses in people. For example, an accent can result in a speaker being perceived as an outsider, someone similar, intelligent, friendly or even untrustworthy (e.g. Nejjari et al., 2012; Purnell, Idsardi, & Baugh, 1999). An L2 accent can also affect the perceived quality and effectiveness of

information transfer (see Nejjari et al. 2012 for an overview), which involves a constant evaluation of the information's quality and the information giver in terms of, among other things, the linguistic features of an utterance and the context in which these are used (Lev-Ari & Keysar, 2010) (Levinson, 1983). This means that, for example, the perceived quality of a lecture taught by an L1 speaker of Dutch with a typically Dutch accent in English can be negatively affected on the basis of their L2 accent in English. Research on L2 accents in a lecturing setting shows that L2-accented speech can negatively impact listeners. For instance, lecturers' L2-accented speech can lead to lower perceptions of status, depending on accent strength, but does not affect a lecturer's likeability (*affect*) (e.g. Dalton-Puffer, Kaltenböck, & Smit, 1997; Hendriks et al., 2016). Furthermore, it can even lead to irritation in students as well as negatively impact perceptions of lecture content understanding (e.g. Bolton & Kuteeva, 2012; Hellekjær, 2010; Hendriks et al., 2016).

The insights on the negative perceptions of status and understanding are valuable, because they show how accentedness affects the judgements listeners have of speakers purely on the basis of accent. However, the reported perceptions of understanding do not reflect actual understanding, as shown by Gerritsen et al. (2010). In their experiment, Gerritsen et al. (2010) had Dutch respondents self-report their understanding of English, with 70 percent indicating they had understood the English; however, only 51 percent was actually able to indicate the correct meaning of the English words used in Dutch.

The definition of understanding employed in Gerritsen et al.'s (2010) corresponds most closely with one of three components of Kachru and Smith's (2008) model of speech understanding: *comprehensibility*, which refers to the degree of understanding the meaning of words individually and in context. Kachru and Smith's model suggests that the process of understanding can be separated into three components. The first component is *intelligibility*, which refers to the manner in which utterances are deciphered into individual sound patterns that form words and sentence-level elements. *Intelligibility* can be measured by requiring listeners to orthographically transcribe individual words or sentences produced by speakers in a recorded setting or by transcribing live voices (see also Nejjari et al., 2012; Nelson, 2011; Yorkston, Strand & Kennedy, 1996). Their second component is *comprehensibility* which is the manner in which words and sentences are understood when it comes to the individual meaning of words and how words put together express meaning within a specific context (see Nejjari et al. 2012). Kachru and Smith (2008) indicate that unlike *intelligibility*, *comprehensibility* also requires that a listener understands the syntax, semantics and even the physical context in which an utterance is heard (Wang 2007;

Yorkston, Strand, & Kennedy, 1996). The third component is interpretability, which is difficult to distinguish from comprehensibility, because both deal with meaning beyond recognition of sound patterns that form words and phrases (Orikasa, 2016). *Interpretability* refers to whether listeners are able to grasp a speaker's intentions and the cultural baggage that is required for discourse strategies to be understood or correctly interpreted (see Nejjari et al., 2012).

To illustrate the three components, for the phrase 'How are you?', *intelligibility* would be high if individual phonemes are heard and combined by listeners to form the individual words and eventually the phrase as a whole. *Comprehensibility* would be high if the listener not only understands the meanings of the individual words and the phrase as a whole, but also understands that this phrase is intended as an enquiry about the listeners' general well-being in an interaction context, for instance, when people run into each other in the street. Finally, *interpretability* would be high if the listener understands that in an Anglo-Saxon cultural context this phrase is commonly used as a phatic greeting that simply requires a greeting in return, and not as an actual inquiry about the listener's life requiring a detailed answer, which is how it is commonly misinterpreted by L2 speakers of English (Kachru & Smith, 2008; Nejjari et al., 2012).

Nejjari et al. (2012) operationalized Kachru and Smith's model of *speech understandability*, to investigate L1 speakers of British English' reactions to a standard British English accent and a strong and a light Dutch English accent. The Dutch English accent, regardless of accent strength, lead to lower *intelligibility* and *comprehensibility*, but did not impact *interpretability*, suggesting that L1 listeners are perfectly capable of understanding the intentions of a speaker with an L2 accent without having heard each word correctly or without perfectly comprehending the content of the information. Nejjari et al.'s study demonstrated that measuring *speech understandability* as a multicomponent construct can yield more nuanced insights into listeners' understanding of L2 accents, in that their results showed that correctly interpreting the intentions of a speaker required only a certain level of *intelligibility* or *comprehensibility*.

As discussed above, L2 accents do not only necessarily affect actual speech understanding, they can also affect the evaluations of speakers or *speaker evaluations*. Accentedness studies show that listeners make personal judgements on the basis of people's accents. In accentedness research, *speaker evaluations* have often been measured in terms of *status* and *affect* (e.g. Zahn & Hopper, 1985; Nejjari et al., 2012; Hendriks et al., 2016; Lambert, Hodgson, Gardner, & Fillenbaum, 1960). A number of

studies have shown that L1 and L2 speakers of English generally ascribe L2 English accents lower *status* compared to L1 English accents, even if presented with speech samples in their own L2 accent in English (e.g. Dalton-Puffer et al., 1997; Matsuura, Chiba & Yamamoto, 1994; McKenzie, 2008; He & Zhang, 2010; Nejjari et al., 2012; Ryan & Sebastian, 1980; Ryan & Bulik, 1982; Cargile, 1997; Cargile & Giles, 1998, Lindemann, 2003). However, L2 accents do not necessarily evoke lower *affect* (Nejjari et al., 2012; Hendriks et al. 2016)

In addition to *status* and *affect*, accentedness research has investigated speaker *dynamism*, albeit less consistently. According to Zahn and Hopper (1985), *dynamism* refers to a person's activity level and enthusiasm. It taps into a listener's perception of the self-presentation of a speaker, and this differs from perceptions of a speaker's *status* or *affect*, as confirmed by van der Haagen (1998). She showed in a study of L2 listeners' perceptions of speakers' *status*, *affect*, and *dynamism* that while a standard British English accent only evoked high *status* and *affect*, a standard American English accent evoked high *status* and *affect* as well as *dynamism*. Grondelaers and van Hout (2015) also showed the relevance of *dynamism* in their study on the perceived prestige of the use of a non-standard object pronoun in Dutch (Dutch: 'hun' (them) in subject position instead of 'hen' (they). Even though the non-standard use of the object pronoun evoked significantly lower *status*, it did not affect the *dynamism* of the speaker, suggesting that, while a speaker who uses this non-standard feature may be perceived as being less cultured, their self-presentation will not be perceived differently from that of a standard language user.

Since the aim of *speaker evaluation* research is to understand listeners' responses to specifically selected accents, it is to know whether listeners are able to identify the accents they are asked to evaluate. Research has shown that a listener's (un) awareness of the language background of a speaker can have an effect on their perceptions of that speaker. For instance, Yook and Lindemann (2013) studied Korean listeners' *status* and *affect* evaluations of different varieties of English. Their results suggest that when an accent is recognized by a listener or made clear to them by researchers, it triggers speaker evaluations connected to a particular group of people (e.g. nationality, culture, sub-culture), and therefore speaker evaluations are connected to ideas or stereotypes listeners have of groups of people. Consequently, in order to study speaker evaluations of specific speaker groups (e.g. Dutch English speakers), we need to know whether listeners are able to correctly identify the speaker's language background.

Finally, most accentedness studies focus on one (professional) communication context, such as: higher education (Bolton & Kuteeva, 2012; Hellekjær, 2010; Hendriks et al., 2016) or business sales (Nejjari et al., 2012; Tsalikis, DeShields & LaTour, 1991). One exception is Cargile (1997) who has investigated L1 and L2 listeners' perceptions of the suitability of Mandarin-accented English in a job interview compared with a higher education classroom. He found that listeners viewed the Mandarin-accented English accent as acceptable in a job interview, but not in a higher education classroom. The acceptability of an accent in one context but not in another could be connected to the Expectancy theory of Burgoon and Burgoon (2001): people have expectations of, among other things, verbal and non-verbal communications that are 'expected' and/or 'desired' in a certain context (2001). By extension, it could be the case that when a speaker has a particular accent in a certain context he or she violates what is an expected or desired accent in that context, which may result in negative perceptions of the speaker or the message on the part of listeners. This suggests that the context in which an accent is heard might impact content understanding and speaker evaluations (cf. Cargile 1997, Dalton-Puffer et al. 1997).

This is why our study attempts to investigate, in a matched-guise experiment, the actual understanding and responses to Dutch-accented English compared to standard British and American English accents in order to find out whether Dutch-accented English hinders the real understanding for L1 speakers of Dutch, and whether this is the case in only a higher education classroom setting or also other professional communication contexts.

4.2 METHOD

4.2.1 Research questions

In order to study the effects of L1 and L2 accents on *speech understandability (intelligibility, comprehensibility, interpretability)* and *speaker evaluations (status, affect, dynamism)*, as well as better understand the impact accent has in several communication contexts, we investigated the reactions of native speakers of Dutch to Dutch-accented English and standard British and American English. Furthermore, we were curious whether the operationalized components of *speech understandability* are correlated with the defined *speaker evaluations* dimensions, and therefore illustrate the potential relationship that may exist between understanding speech and evaluating speakers, which apart from Nejjari et al. (2012) has not been studied before.

To this aim, three research questions were formulated:

- RQ1: Do *accent* (standard British English, standard American English, Dutch English) and *context* (lecture; audio tour; job pitch) affect Dutch listeners' *speech understandability (intelligibility, comprehensibility, interpretability)*?
- RQ2: Do *accent* (standard British English, standard American English, Dutch English) and *context* (lecture; audio tour; job pitch) affect Dutch listeners' *speaker evaluations (status, affect, dynamism)*?
- RQ3: Are *speech understandability* and *speaker evaluations* correlated?

4.2.2 Speakers: matched-guise speaker, control, and filler speakers

In our experiment, a matched-guise speaker was selected to produce the three accents in order to avoid responses to the voice characteristics of individual speakers. The matched-guise speaker had been assessed in an earlier accentedness experiment, which showed that he could produce the three accents under study that represented the independent variable *accent*: (1) standard British English, (2) standard American English, and (3) the typical English accent of L1 speakers of Dutch (Nejjari et al., 2019). In the current study, the standard accents of British and American English refer to accents generally associated with the national accent norm of these nations. We regard a typical Dutch English accent in the present study as containing features that L1 speakers of Dutch and others familiar with Dutch and Dutch English will recognize as such. For example, because Dutch lacks dental consonants [ð] as in *this, mother, breathe* and [θ] as in *think, Martha, breath,* they are often mispronounced as stop consonants, [d] and [t] respectively, by Dutch speakers of English. Dutch also lacks voiced fricatives and plosives in the coda, causing the voiced obstruents of English to generally be pronounced as their voiceless counterparts in Dutch speakers' English (e.g. *live, badge, bad, bag* will be said with [f, tʃ, t, k]) (Gussenhoven & Broeders, 1997). As no standard has been defined for a standard Dutch English accent, in this study it is defined as being 'typical' (Nejjari et al., 2019).

To prevent listeners from deducing that they were listening to the same matched-guise speaker a number of times, we included (speech samples from) six male control speakers as stimuli: two L1 speakers of standard British English, two L1 speakers of standard American English, two L1 speakers of Dutch who have a typical Dutch accent in English. All but one of these speakers had been assessed on the representativeness of their accents in the same accentedness experiment as the matched-guise speaker (see Nejjari et al., 2019). The speaker who was not assessed was one of the Dutch-accented English control speakers. However, he was generally regarded by experienced linguists and Dutch language specialists as a representative speaker of

Dutch-accented English. In addition, one further male speaker produced a speech sample that was presented to listeners at the beginning of the experiment (the filler speech sample) to familiarize them with the task at hand. The filler speaker had also been assessed in the earlier accentedness experiment and was accepted, by L1 speakers of British English, as an L1 and standard speaker of British English (Nejjari et al., 2019).

4.2.3 Instrumentation and participants

Participants were asked to evaluate four speech samples by four different speakers (see section *4.5.1 Speech sample links*). They were first required to evaluate the filler sample followed by the three samples (in this order: Lecture context, the Audio Tour context, Job Pitch context) produced by the matched-guise speaker and the control L1 speakers. To ensure that the nine (plus 1 filler) matched-guise samples could be evaluated in each context for each accent, to avoid repeating the content of each context, and to limit any order effect, 18 listener groups were created, with approximately 30 listeners per group.

Our participants were highly educated (12.3% A-level; 51.2% bachelor; 27.6% master; 2.3% PhD; 6.3% other) native speakers of Dutch (mean age 39; 60% female, 40% male) with no background in linguistics. This listener group was selected because they represent the part of Dutch society most likely to use English in an educational and professional setting (Bouma, 2016; Lizzini et al., 2017). Participants were asked to indicate what they believed the speakers' country of origin was. Over 92% correctly indicated that the Dutch English matched-guise speech sample was by a speaker that was originally from the Netherlands. For the British English speech samples 88% correctly indicated that the speech sample was by a speaker that was originally from Great Britain, and over 81% correctly indicated that the standard American English matched-guise speech sample was by a speaker that was originally from the United States.

Only the responses of listeners who had correctly identified the speakers' country of origin were used in the analyses. Listeners who were not able to correctly identify speakers' country of origin were excluded from the analyses, as their assessments could reflect interfering associations with other speaker groups, and thus with other accents than those under study. Furthermore, to understand whether self-reported English fluency impacted *speech understandability*, the listeners were asked to estimate their English language skills (listening, reading, speaking, writing) to see whether this would impact their responses on a 5-point Likert scale (1: very low; 2: low; 3: average; 4: high; 5: like a native speaker). Seventy-three percent (N=392) indicated that their skills for all four areas were average or higher, and these

responses were analyzed to answer the main research questions, except for the correlations between the *speech understandability* dimensions and the correlations between the listeners' estimated English language skills and their *speech understandability.* These correlations were calculated using all responses (N=545), instead of only the responses by listeners with at least average English language skills (N=392) to be able to provide a more complete assessment of the relationship between the *speech understandability* dimensions and how these might be connected with estimated English language skills.

4.2.4 Stimuli

One filler text (on a general topic) and three texts that represented the independent variable *context*: (1) an introduction to a marketing lecture; (2) an art gallery audio tour segment; (3) a job pitch for a retail management position were used as the basis for the speech samples (see section *4.5.1 Speech sample links*). All but the filler text reflect three main contexts in which English is an important lingua franca: higher education, tourism, and international business (Gerritsen et al., 2016). The lecture, audio tour, and filler texts were selected from an IELTS Academic English test and the job pitch text was taken from a human resources webpage. The matched-guise speaker produced the three accents in all three contexts, resulting in nine speech samples. The six control speakers produced English in their L1 accents (Dutch English, standard British and American English) in the three contexts, resulting in 18 speech samples, and the additional speaker produced one filler speech sample on a general topic in standard British English.

4.2.5 Speech understandability

Following Kachru and Smith (2008) and Nejjari et al. (2012), three questions were used to measure *speech understandability* and more specifically the ability (1) to literally recognize words (*intelligibility*); (2) to understand the meaning of the words within the context (*comprehensibility*), and (3) to understand the intention of the speaker/purpose of the message (*interpretability*).

To measure *intelligibility*, listeners were presented with a speech sample consisting of the first 9-12 words of each of the four stimuli speech sample texts and asked to write down what was literally stated. We counted the number of words correctly transcribed, and an Intraclass Correlation Coefficient for two raters turned out to be extremely high (.98). To measure *comprehensibility* listeners were asked to indicate whether one statement per speech sample on the topic of the that particular speech sample was correct or not. Finally, *interpretability* was measured by having

listeners indicate whether one statement per speech sample on the communicative intentions of the speaker was correct or not.

4.2.6 Speaker evaluations

To assess *speaker evaluations*, the listeners were asked to indicate on 5-point Likert scales (1 = strongly disagree; 5 = strongly agree; 3 = neither disagree nor agree) to which extent they believed the speaker possessed 11 personality traits. These traits represent the three dimensions of *speaker evaluations* this experiment focuses on, namely *status* (competent, educated, having authority, intelligent and cultured), *affect* (considerate, pleasant and friendly), and *dynamism* (energetic, enthusiastic, confident). The traits associated with *status* and *affect* are based on Nejjari et al. (2012). *Status* represented the degree to which a speaker was viewed as being intelligent and well-educated, and the *affect* represented the degree to which a speaker was perceived as being likeable. *Dynamism* measured the self-presentation of the speaker and was based on Grondelaers, van Gent et al. (2015). A factor analysis, using a principle axis factoring extraction method with an Eigenvalue >1 criterion for factor extraction, followed by a varimax rotation, on the ratings of the personality traits showed a resolution into three factors: *status, affect, dynamism*.

Table 1. Rotated Factor Matrix: factor loadings of the scores on 11 scales with three factors. Only loadings >.550 have been printed.

	Factor 1	**Factor 2**	**Factor 3**
	Status	Affect	Dynamism
Competent	.690		
Considerate		.805	
Cultured	.855		
Highly educated	.871		
Pleasant		.751	
Energetic			.766
Authoritative	.596		
Friendly		.868	
Enthusiastic			.670
Intelligent	.855		
Confident			.671

4.2.7 Procedures

An online questionnaire was conducted via a student Facebook page (https://www.facebook.com/questionnairestudent.nl) and an online data collection service. On average, listeners needed 15 minutes to complete the questionnaire. Approximately 41% of the original data was discarded which resulted in data from 545 listeners. The reasons for excluding data included: filling in nonsense answers or numbers, symbols; not speaking the required L1 language; only providing neutral answers to all questions in the questionnaire.

4.3 RESULTS

This study's findings will be reported for the three RQs successively. All mean and frequency measurements of the *speech understandability* and the *speakers evaluation* of the three accents produced in three contexts are presented in Table 2 (see section 4.5.2).

4.3.1 Speech understandability, accent, context (RQ1)

In general, *speech understandability* was high: *comprehensibility* was 89.2%, *interpretability* was 81.0%. The mean correctly transcribed words (*intelligibility*) was 9.23, which is fairly high given the maximum possible correct score of 11 (Lecture, Job Pitch) or 12 (Audio Tour) words. The results for *intelligibility* in relation to *context* and *accent* are shown in Figure 1, with 95% confidence intervals. The results show substantial overlap in the confidence intervals between the bars, indicating that there are no strong differences. An analysis of variance was applied to investigate the effects of *context* and *accent* and their interaction. The interaction effect was significant ($F(4, 383)=2.84$, $p=.02$, partial $\eta^2 = .03$). There was a significant main effect for *accent* ($F(2, 383)=4.30$, $p=.01$, partial $\eta^2 =.02$), but not for *context* ($F(2, 383)=1.27$, $p=.28$, partial $\eta^2 =.01$). However, the post-hoc test (HSD) for *accent* did not show significant differences.

The results for *comprehensibility* in relation to *context* and *accent* are shown in Figure 2, with 95% confidence intervals. There are no error bars for the Lecture in British and American English because the comprehension scores were 100%. The remaining error bars results showed substantial overlap in the confidence intervals between the bars, indicating that there are no strong differences between the three accents and the three contexts. A logistic regression was applied to test the effects

Figure 1. Mean proportions of correct intelligibility for accent (British English, American English, Dutch English) and context (Audio Tour (Min=0; Max=12 words); Lecture (Min=0; Max=11 words); Job Pitch (Min=0; Max=11 words))

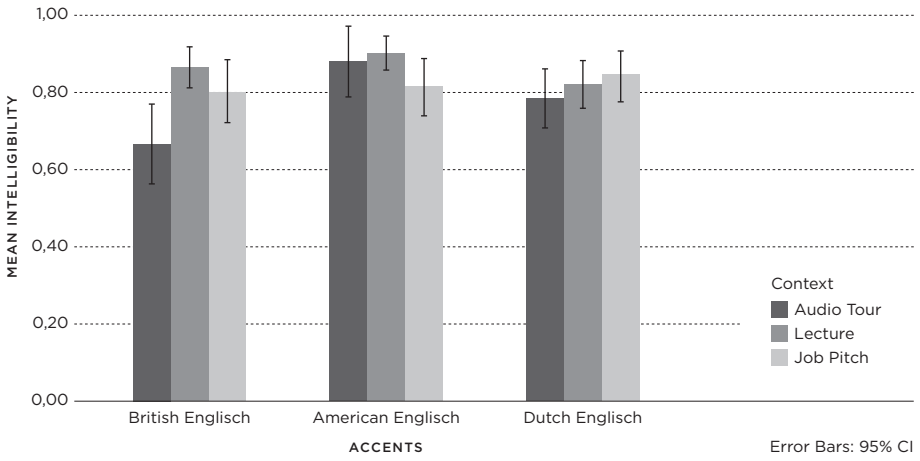
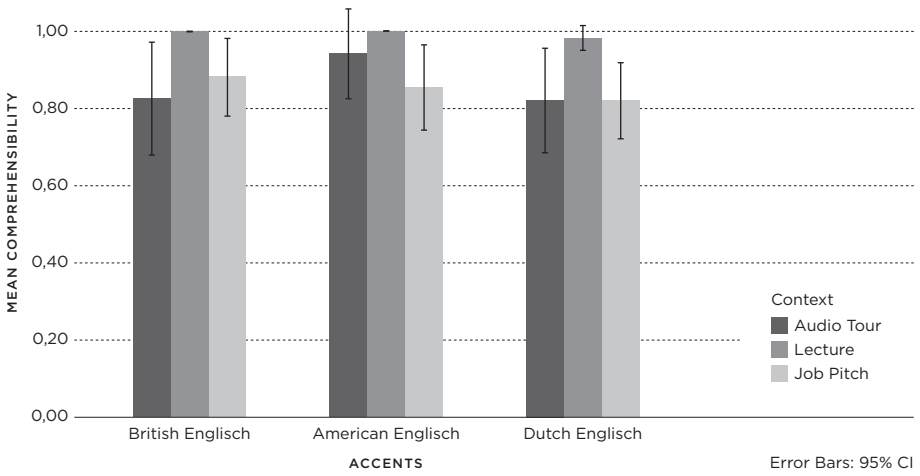


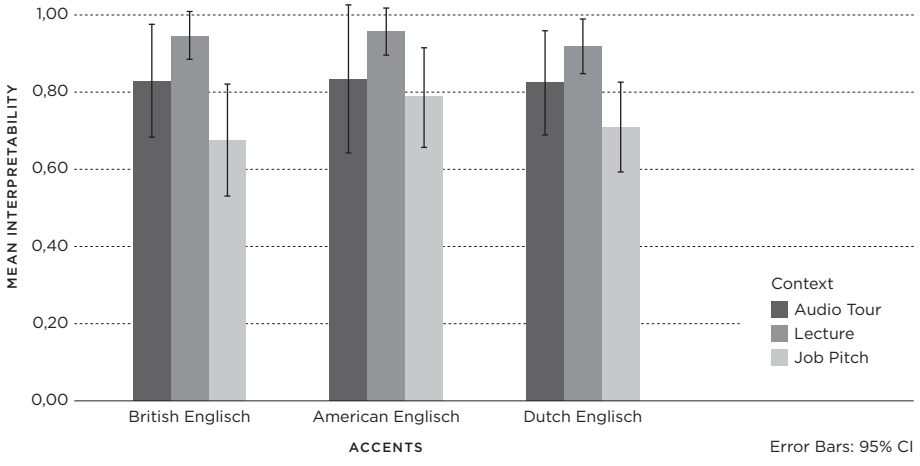
Figure 2. Mean proportions of correct comprehensibility for accent (British English, American English, Dutch English) and context (Audio Tour, Lecture, Job Pitch).



of both *context* and *accent* as well as their interaction. The interaction effect could be removed (Deviance score=5.58, df=4, p=.23). The same applied to accent in the next step (Deviance score=1.80, df=2, p=.41). The remaining factor context was significant (Wald=17.77, df=2, p=.00). Post-hoc analysis (Bonferroni) revealed that the *comprehensibility* of the Lecture was significantly higher compared to the Audio Tour and the Job Pitch.

The results for *interpretability* in relation to *context* and *accent* are shown in Figure 3, with 95% confidence intervals. The results demonstrated substantial overlap in the confidence intervals between the bars, indicating that there are no strong differences between the three accents and three contexts. A logistic regression was applied to test the effects of both *context* and *accent* as well as their interaction. The interaction effect could be removed (Deviance score between the model with the interaction and the model without interaction was 3.00, df=4, p=.56). The same applied to *accent* in the next step (Deviance score=.42, df=2, p >.05). The remaining factor *context* was significant (Wald=33.85, df=2, p=.00). Post-hoc analysis (Bonferroni) revealed that *interpretability* was significantly higher for the Lecture compared to the Audio Tour and the Job Pitch.

Figure 3. Mean proportions of correct interpretability for accent (British English, American English, Dutch English) and context (Audio Tour, Lecture, Job Pitch.



The results show that there are ceiling effects that could have had a suppressing effect on the correlation between the understandability measures. *Interpretability* showed a significantly positive, but weak, correlation with both *comprehensibility* (r(545)=.26, p=.00) and *intelligibility* (r(545)=.20, p=.00). *Comprehensibility* showed a significant, but again weak, positive correlation with *intelligibility* (r(545)=.13, p=.00). As explained in the method section, the participants were asked to report their estimated English language skills. In order to check whether their estimated language skills had had an effect on *speech understandability*, correlations were computed with the three understandability variables. None of the correlations were significant: *interpretability*: r(544)=.07, p=.11; *comprehensibility*: r(544)=-.02, p=.70; intelligibility: r(545)=.15, p=.06.

4.3.2 Speaker evaluations, accent, context (RQ2)

The results for status in relation to *context* and *accent* are shown in Figure 4, with 95% confidence intervals. An analysis of variance was applied to investigate the effects of *context* and *accent* and their interaction. The interaction effect was not significant (F(4, 383)=.53, p=.72, partial η^2=.01). There was a significant main effect for accent (F(2, 383)=106.86, p=.00, partial η^2=.36), but not for *context* (F(2, 383)=2.15,

Figure 4. Mean scores of status (1=negative; 3=neutral; 5=positive) for accent (British English, American English, Dutch English) and context (Audio Tour, Lecture, Job Pitch).

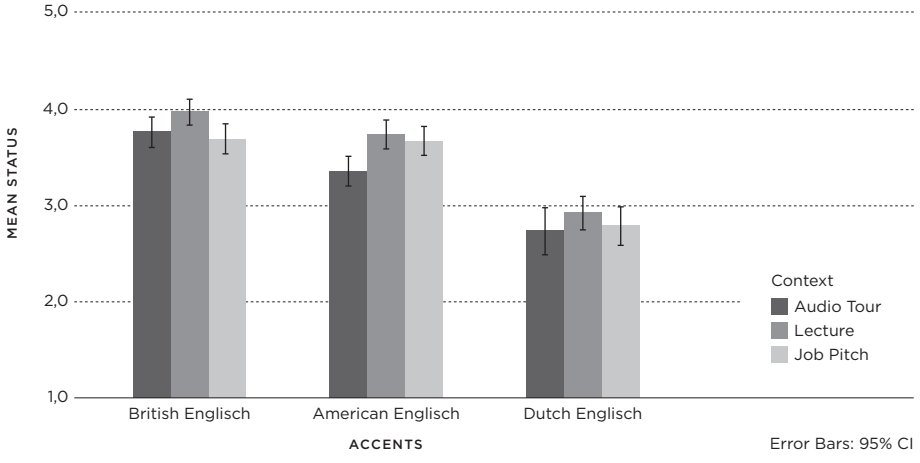
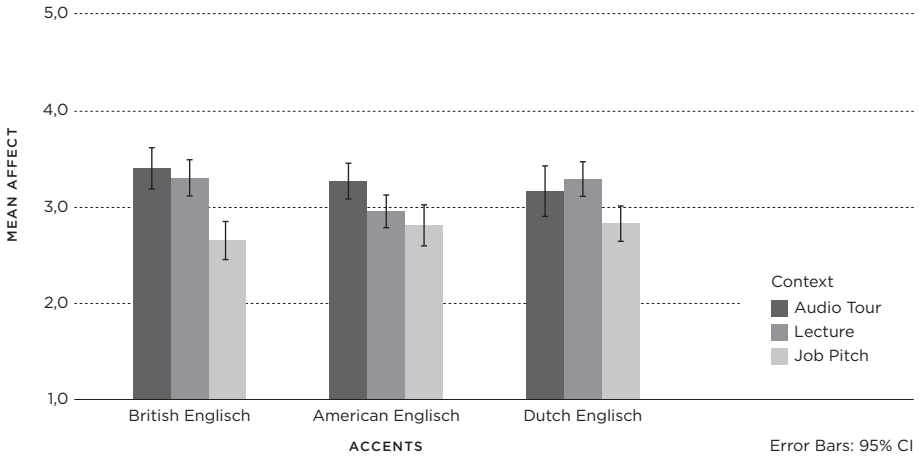


Figure 5. Mean scores of affect (1=negative; 3=neutral; 5=positive) for accent (British English, American English, Dutch English) and context (Audio Tour, Lecture, Job Pitch).

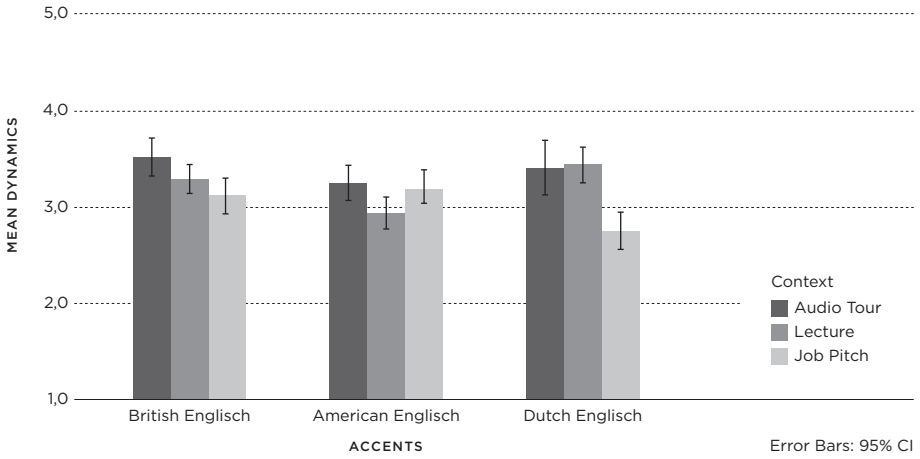


p=.12, partial η^2=.01). Post-hoc comparisons (HSD) showed significantly higher *status* for British English and American English compared to Dutch English (see section 4.5.2).

The results for *affect* in relation to *context* and accent are shown in Figure 5, with 95% confidence intervals. An analysis of variance was applied to investigate the effects of *context* and *accent* and their interaction. The interaction effect was not significant (F(4, 381)=1.90, p=.11, partial η^2=.02). There was no significant main effect for *accent* (F(2, 381)=1.07, p=.31, partial η^2=.01), but *context* was significant (F(2, 381) =20.29, p=.00, partial η^2=.10). Post-hoc comparisons (HSD) showed that the Audio Tour and Lecture aroused significantly higher *affect* than the Job Pitch (see section 4.5.2).

The results for *dynamism* in relation to *context* and *accent* are shown in Figure 6, with 95% confidence intervals. An analysis of variance was applied to investigate the effects of *context* and *accent* and their interaction. The interaction effect was

Figure 6. Mean scores of dynamism (1=negative, 3=neutral, 5=positive) for accent (British English, American English, Dutch English) and context (Audio Tour, Lecture, Job Pitch).



significant (F(4, 384)=75.59, p=.00, partial η^2=.06). There was no significant main effect for *accent* (F(2, 384)=1.42, p=.24, partial η^2=.01), but *context* was significant (F(2, 384)=8.19, p=.00, partial η^2=.04). Post-hoc comparisons (HSD) showed that Job Pitch evoked significantly lower *dynamism* than the Audio Tour and Lecture (see section 4.5.2), but the significant interaction clarified that this was due to the negative evaluation of the Dutch English Job Pitch.

4.3.3 Correlation speech understandability and speaker evaluations (RQ3)

The correlations between *speech understandability* and *status* were not significant (*interpretability*: r(391)=.08, p=.10; *comprehensibility*: r(392)=.05, p=.33; *intelligibility*: r(392)=.02, p=.68). The correlations between *speech understandability* and *affect* were weak and negative only for *comprehensibility* (*interpretability*: r(389)=.01, p=.99; *comprehensibility*: r(390)=-.110, p=.03; *intelligibility*: r(390)=-.07, p=.17). The correlations between *speech understandability* and *dynamism* were not significant (*interpretability*: r(392)=.03, p=.59; *comprehensibility*: r(393)=-.04, p=.44; *intelligibility*: r(393)=-.09, p=.09).

4.4 CONCLUSION AND DISCUSSION

Our three research questions investigated the effect *accent* and *context* had on *speech understandability* (RQ1), *speaker evaluations* (RQ2), and whether *speech understandability* and *speaker evaluations* were correlated (RQ3).

4.4.1 Speech understandability, accent, context (RQ1)

The effects of *accent* and *context* on *speech understandability*(RQ1) are only significant for *context*. *Accent* did not have any effect on the three measures of *speech understandability*. This indicates that Dutch-accented English is as comprehensible as standard British and American English.

Context had an effect on *interpretability* and *comprehensibility*. The Lecture context was more interpretable and comprehensible than the Audio Tour context and the Job Pitch context regardless of *accent*. These findings may be explained by the listeners' backgrounds. Most were highly educated, and might have had more experience with an academic educational setting than with an audio tour or a job pitch. These results are in line with Cargile (1997), who found that if listeners are confronted with an accent in a context they did not expect, this might affect their reactions to a speaker.

It should be noted that *speech understandability* was high for all accents and contexts for all three *speech understandability* components (section 4.5.2). This caused a ceiling effect but also contributed to weak, positive correlations between the three dimensions. This means that when listeners were able to decipher sound patterns into words and phrases (*intelligibility*), this positively influenced their ability to understand these words and phrases (*comprehensibility*) and understand a speaker's intention (*interpretability*). What needs to be noted is that *comprehensibility* and *interpretability* were measured by asking listeners to confirm or deny the validity of a statement on the content (*comprehensibility*) and communicative purpose (*interpretability*) of the sample. This research method might have positively impacted the found results; however, the positive correlation between the three components does suggest that higher *intelligibility* helps other components of understanding.

The experiment was conducted with highly educated Dutch listeners (at least 80% of listeners had obtained bachelor level education), who are generally regarded as having good English language skills (EF, 2018). This could mean that the listeners had a high enough English fluency, to easily deal with the content from academic and professional sources that they were presented with in the experiment. The listeners selected for this experiment were those people who – in the Netherlands – would

most likely be confronted with all three contexts precisely because they are highly educated and most likely to encounter English in university lectures, in an audio tour in a museum or art gallery, and in interactions on the international job market.

4.4.2 Speaker evaluations, accent, and context (RQ2)

With regard to the effects of accent and context on *speaker evaluations* (RQ2), it can be concluded that having a Dutch English accent, has a significantly negative effect on a speaker's *status* compared with standard British and American English. *Context* had no significant effect on the perceived *status* of a speaker. This indicates that *accent* matters more in perceptions of the *status* of a speaker than *context*. In terms of perceptions of *affect* and *dynamism*, *context* appears to be more important than *accent*, because *context* had a significant effect on the *affect* and *dynamism* ascribed to the speakers, with the Job Pitch context evoking significantly lower *affect* and *dynamism* compared to the Lecture and Audio Tour contexts. Yet, even though the means for all three contexts were mostly neutral, and despite the fact that for *dynamism* and *context* the statistically significant results were caused by the Dutch English job pitch, it is striking that the Job Pitch context evoked such a negative response.

The above results concur with previous research that found that L1 and L2 speakers of English tend to assign L2 English accents lower status compared to L1 English accents, including their own L2 accent in English (e.g. Nejjari et al., 2012; Cargile, 1997; Cargile & Giles, 1998; Dalton-Puffer et al., 1997; He & Zhang, 2010; Hendriks et al., 2016; Lindemann, 2003; Matsuura et al., 1994; McKenzie, 2008; Ryan & Sebastian, 1980; Ryan & Bulik, 1982). Similar to other research, our results also showed that an L2 accent does not necessarily have to evoke lower *affect* (Nejjari et al., 2012; Hendriks et al., 2016) or *dynamism* (Nejjari et al., to be published), and that *speaker evaluations* can be impacted by *context* in a significant manner (Cargile, 1997).

4.4.3 The correlation between speech understandability and speaker evaluations (RQ3)

The last research question focused on the correlations between *speech understandability* and *speaker evaluations* (RQ3). It was established that *status* and *dynamism* were not significantly correlated with *speech understandability*. This means that a listener's evaluation of the *status* and *dynamism* of a speaker is not related to their ability to comprehend a speaker. For *affect* there was a weak, negative correlation with *comprehensibility*, meaning that the lower the *comprehensibility* of the speaker, the higher the *affect* towards the speaker. These correlation results contradict

Nejjari et al.'s (2012) results that showed a significantly positive correlation between *status* and *intelligibility* and *comprehensibility* as well as between *affect* and *comprehensibility*. The reason for this may be that in their study other listeners were involved: British listeners who reacted to standard British English accent and two levels of Dutch-accented English. Half of the listeners were not familiar with Dutch accented-English. As a result, the listeners would have probably needed to put more effort into comprehending the accents and the speech content than the L1 speakers of Dutch in the current study, who were all familiar with all three accents, and achieved extremely high levels of *speech understandability* which resulted in the ceiling effects discussed above.

4.4.4 Limitations and future research

This study has a number of limitations. First, we used only one male matched-guise speaker in our experiment. This could have led to reactions that represent a response to male speakers only, and to this particular male. Second, the listeners were all Dutch and we therefore do not know how other listeners with other language backgrounds would respond to the three accents. Finally, the listeners were selected to represent people who in general would be most likely to be familiar with the three selected contexts. With hindsight, however, professionals who regularly interact in these contexts, such as HR managers in the job interview context, would have been the optimal choice, but this was not possible for practical reasons.

The significant results for *context* in this study provides an important nuance to L1 and L2 accentedness studies. Future research should include other contexts and more L2 and L1 accents. This can help language learners understand the effects their accents and provide insights into which accents are deemed desirable for which contexts, and can lay bare speaker perceptions that may exist within different listener groups. Furthermore, the assessment of *speech understandability* in terms of Kachru and Smith's (2008) three components: *interpretability*, *comprehensibility*, and *interpretability*, yielded results that provide useful insights into the levels at which listeners are able to understand speech. However, the research method employed to measure *comprehensibility* and *interpretability*: multiple-choice questions on the content and purpose of the speech samples, resulted in high scores and ceiling effects which could suggest that the questions might have positively impacted the results. Therefore, this research method should be further validated with different listener groups.

The use of the matched-guise technique contributed to the validity of our results, since they cannot be attributed to the voice characteristics of individual speakers

(as might have been the case had verbal guises been used). Although our methodology yielded insightful results, future research should assess how L2 and L1 English accents affect L1 and L2 English speakers' behavior in specific contexts, similar to Purnell, Idsardi and Baugh's study (1999). With the increasing globalization of not only academia, but also the international job market it is of great importance to research in these particular communication contexts how people view one another and how they behave towards each other on the basis of accents.

4.4.5 Implications

Accent training aimed at becoming as L1 as possible for Dutch learners of English should not be emphasized in language training, if the intent is to be understandable, and evoke higher affect and dynamism. However, status is negatively affected by L2 accentedness in English, and it therefore pays to sound L1 if the intent is to evoke perceptions of high status. In our study context has also proven to have a significant impact on attitudes towards speakers and creating awareness of such potential effects can help learners understand the impression they make in their communications in English.

4.5 SUPPORTING INFORMATION

4.5.1 Speech sample links study 3 (and 4): matched guises, filler, controls

The speech sample links below redirect to one speech sample per accent (standard British, standard American English, Dutch English) and communication context (lecture, audio tour, job pitch), per speaker (matched-guise speaker; filler speaker; control speakers). For the controls, three speech sample links are available as examples (for each accent one).

Matched-guise speaker Lecture

- standard British English https://cls.ru.nl/webexp-media/HV_BE_L.html
- standard American English https://cls.ru.nl/webexp-media/HV_AE_L.html
- Dutch English https://cls.ru.nl/webexp-media/HV_DE_L.html

Matched-guise speaker Audio tour

- standard British English https://cls.ru.nl/webexp-media/HV_BE_A.html
- standard American English https://cls.ru.nl/webexp-media/HV_AE_A.html
- Dutch English https://cls.ru.nl/webexp-media/HV_DE_A.html

Matched-guise speaker Job pitch

- standard British English https://cls.ru.nl/webexp-media/HV_BE_J.html
- standard American English https://cls.ru.nl/webexp-media/HV_AE_J.html
- Dutch English https://cls.ru.nl/webexp-media/HV_DE_J.html

Filler speaker

- standard British English https://cls.ru.nl/webexp-media/CG_BE_F.html

Examples of controls or native speakers for each accent (in the lecture context only)

- standard British English https://cls.ru.nl/webexp-media/JC_BE_L.html
- standard American English https://cls.ru.nl/webexp-media/PG_AE_L.html
- Dutch English https://cls.ru.nl/webexp-media/CJ_DE_L.html

4.5.2 Results speech understandability and speaker evaluations

Table 2. Mean, % speech understandability; mean speaker evaluations, per accent (British English, American English, Dutch English; 1=negative; 3=neutral; 5=positive; 95% confidence interval) and context (Lecture, Audio Tour, Job Pitch).

	Speech understandability			**Speaker evaluations**		
Accent, Context	Intelligibility Mean (SD)	Comprehensibility % correct	Interpretability % correct	Status Mean (SD)	Affect Mean (SD)	Dynamism Mean (SD)
Dutch English Lecture*	9.02(2.69) n=63	95.8% n=68	88.6% n=62	2.84(.70) n=63	3.28(.75) n=62	3.43(.76) n=63
Dutch English Audio Tour*	9.41(2.64) n=34	82.9% n=29	82.9% n=29	2.69(.70) n=39	3.10(.67) n=32	3.36(.81) n=34
Dutch English Job Pitch**	9.26(2.79) n=62	78.8% n=52	71.2% n=47	2.76(.78) n=62	2.80(.75) n=61	2.72(.76) n=62
British English Lecture	9.49(2.16) n=55	100% n=69	94.2% n=65	3.99(.57) n=55	3.34(.78) n=55	3.30(.60) n=55
British English Audio Tour	7.97(3.24) n=29	82.5% n=52	84.1% n=53	3.85(.64) n=29	3.40(.74) n=29	3.49(.71) n=29
British English Job Pitch	8.84(2.79) n=43	86.8% n=59	67.6% n=46	3.74(.58) n=43	2.68(.80) n=43	3.14(.74) n=43
American English Lecture	9.91(1.63) n=47	100% n=61	93.4% n=57	3.74(.55) n=46	2.94(.68) n=47	2.96(.72) n=47
American English Audio Tour	10.56(2.18) n=18	80.3% n=53	86.4% n=57	3.53(.36) n=18	3.33(.54) n=18	3.39(.57) n=18
American English Job Pitch	8.95(2.61) n=42	82.1% n=55	70.1% n=47	3.71(.65) n=42	2.67(.84) n=42	3.17(.76) n=42

*=max. 11 words; ** max. 12 words intelligible. N=392; n=number of listeners per accent and context.

5
WHERE DOES A ‘FOREIGN’ ACCENT MATTER?

ABSTRACT

How well L2 English is understood and how L2 English speakers perceive one another within varying communication contexts has been studied relatively rarely, even though most speakers of English in the world are L2 speakers. In this matched-guise experiment (N=1699) the effects of L1 and L2 English accents and communication context were tested on speech understandability *(intelligibility, comprehensibility, interpretability)* and speaker evaluations (status, affect, dynamism). German (N=617), Spanish (N=540), and Singaporean listeners (N=542) were asked to evaluate three accents (Dutch-accented English, standard British English, standard American English) in three communication contexts (Lecture, Audio Tour, Job Pitch). The main findings are that a Dutch-accented English accent was understood equally well as the two L1 English accents. Furthermore, Dutch-accented English evoked equally positive evaluations compared to L1-English accents from German listeners, and for Spanish and Singaporean listeners evoked even more positive evaluations than the two L1 English accents. These results suggest that for English language learners additional accent training may not always be necessary, especially when they mostly interact with other L2-speakers of English. More generally, our results indicate that L2 English speakers' understanding and their evaluation of L1 and L2 Englishes do not reflect traditional L1 norms. Instead, our results illustrate the socio-cultural embedding of a language norm in an LFE speech community that does not view accent varieties as hindrances to successful communication.

This chapter is based on:
Nejjari, W., Gerritsen, M., Planken, B. & van Hout, R.
(Submitted and revision in progress). Where does a 'foreign' accent matter? German, Spanish and Singaporean listeners' reactions to Dutch-accented English compared to standard British and American English accents in three communication contexts.

5.1 INTRODUCTION

The widespread use of English in international interactions has led to discussions on how varieties of native (L1) and non-native or 'foreign' (L2) English are understood and perceived by native and non-native English listeners (e.g. Bolton & Kuteeva, 2012, Hellekjaer, 2010; Dalton-Puffer et al., 1997; Hendriks et al., 2016; Chiba, Matsuura, & Yamamoto, 1995; McKenzie, 2008; Nejjari et al. 2012; Nejjari et al. to be published; He & Zhang, 2010; Ryan & Sebastian, 1980; Ryan & Bulik, 1982; Cargile 1998; Lindemann, 2002). This also holds true for the Netherlands, where English is the most dominant L2 language in education, advertising, academia, politics, and business and the most important lingua franca in transnational communications (e.g. Edwards, 2016; Nickerson, 2005). The Dutch often criticize the English spoken by their fellow Dutchmen, because it is thought to hinder effective communication, for example in the international classroom (e.g. Bouma, 2016; Bronkhorst, 2015; van Gaal, 2018; Huygen, 2017). Such assumptions have to some extent been confirmed in an experiment by Hendriks, van Meurs and Hogervorst (2016), who found Dutch students' evaluations to be significantly more negative with regard to speech understanding when a lecture was taught by a lecturer with a strong Dutch English accent compared to when a lecture was taught in a slight Dutch English accent or in Dutch. These views are common, despite the fact that, generally, Dutch speakers of English are considered proficient L2 speakers of English (e.g. Edwards, 2016; ranked nr. 2 out of 88 nations according to EF, 2018), and some have even argued that English has integrated into Dutch society to such an extent that it is gaining a status similar to the English spoken in former British colonies, such as Singapore, Nigeria, and India (see Gerritsen et al. 2016).

5.1.1 Second Language Acquisition (SLA) versus Lingua Franca English (LFE)

The negative perceptions regarding the impact of Dutch-accentedness in English on effective speech understanding and information transfer might reflect a general Dutch societal language norm based on the view that proficiency as an L2 English speaker is only truly achieved when a speaker's language skills match those of an L1 English speaker, in terms of for instance pronunciation. This viewpoint mirrors the traditional SLA research perspective on languages and speakers, which is based on the notion that languages are used by 'native speakers' (L1) and learned by 'non-native speakers' (L2), and that acquiring an L2 involves achieving proficiency in a 'target language', free from L1 influence (e.g. 'interlanguage') (e.g. Corder, 1967; Selinker, 1972; Richards, 1974; Ellis, 1994; Gass, 1994; Lightbown & Spada, 1999; Penning de

Vries, Cucchiarini, Strik & van Hout, 2019). This perspective has been debated by many and criticized by some for assuming that L2 speakers are at a linguistic disadvantage because they are linguistically and communicatively less competent compared to L1 speakers (e.g. Canagarajah, 2007; Firth & Wagner, 1997; Kachru, 1982). The focus of traditional SLA appears to not take into consideration the impact of language use on, for instance, the evaluations of speakers' competencies and likeability, and how language use affects a listeners' evaluation of a speaker's self-presentation. More generally then, traditional SLA does not sufficiently take the social and interactive nature of language use into consideration. With respect to the latter, for example, Firth and Wagner (Firth & Wagner, 1997) called for a more holistic view on SLA and argued the importance of acquiring more knowledge on the social context, the interactive nature of language acquisition, and the effects of language use on the perceptions of speakers, since this might lead to more insight into the contexts and practical circumstances in which language is used and learned, and when interactions are successful.

Canagarajah (2007) views the position of English as a global lingua franca as a prime example of linguistic circumstances that illustrate that the traditional SLA perspective is not optimal. Most speakers of English in the world are L2 speakers with various cultural and linguistic backgrounds (Crystal, 2009) who communicate with one another in a variety of societal domains, such as business, politics, and academia. Canagarajah suggests that speakers who use English as a global lingua franca (i.e. lingua franca English or LFE) have created a worldwide speech community, that is, an international group of L2 English speakers, not separated by traditional national and linguistic boundaries, but view English as a resourceful tool to achieve their objectives (Kankaanranta & Planken, 2010). At the same time, individual L2 English speakers in this community are members of their own separate linguistic communities, for example within nation states. This means that so-called 'LFE speech community members' can have multiple linguistic identities, and that linguistic heterogeneity and flexibility are the norm within the LFE speech community. Consequently, LFE speech community members might not be focused on how fellow L2 English speakers are dissimilar to L1 English speakers and how that may potentially hinder communications and reflect (potentially negatively) on how they are perceived as speakers.

Canagarajah's assumption of the existence of an LFE speech community and its members' potential flexible language attitudes is supported by some speech evaluation research that has shown that L2 English language users are more tolerant of different L2 English varieties than of L1 English varieties (e.g. Bernaisch, 2012). On

the other hand, negative evaluations of L2 English compared to L1 varieties, or in general, have also been observed (Ahn, 2014; He & Miller, 2011; McDonald & McRae, 2010; McKenzie, 2008; Rivers, 2011; Sasayama, 2013). In order for learners, but also teachers, to navigate this new linguistic reality of an LFE speech community in which multiple varieties of English exist and are evolving, more knowledge is required on the responses to different varieties of English. The current study takes Dutch-accented English as an example of L2 English to assess how global LFE speech community members with different L1 backgrounds (than Dutch) understand and perceive L2 English compared to varieties of L1-accented English (British and American).

5.1.2 Dutch English in the LFE speech community: navigating heterogeneous linguistic spaces

Canagarajah (2007) argues that L2 English is used in interactions between speech community members with a variety of cultural and linguistic identities. Therefore, it is not only relevant from an SLA research perspective to determine the effects of L2 English use, but also from a pragmatic viewpoint, as it is important in different professional settings for speakers to understand how their L2 English use affects their communications with specific speaker groups in the larger LFE community. For example, Dutch L2 English speakers are likely to interact with fellow LFE speech community members from nations that are important to them in terms of trade, politics, and academia. Germany, which neighbors the Netherlands, is its most important trade partner (RVO 'Netherlands Enterprise Agency', 2019a; Central Bureau for Statistics, Netherlands, 2019). Like Dutch, Germany's only official language, German, is a West-Germanic language, and English is spoken in Germany at a high level (ranked in the highest high English fluency nations: nr. 10 out of 88 nations according to EF, 2018). This means that the linguistic distance between Dutch and German L2 English speakers might be smaller compared to, for instance, the distance between Dutch and Spanish L2 English speakers, another important trade partner to the Netherlands (RVO, 2019b, Greven, 2016; Gijsberts, Lubbers, Fleischmann, Maliepaard & Schmeets, 2016). Spain has Spanish, a Romance language, as its official language (next to other official, regional languages: Basque, Catalan and Galician), which is relatively more distant (than German) from Dutch. In addition, Spain does not have a high English fluency level (ranked 34 out of 88 nations tested, EF, 2018).

In addition to possessing lower English fluency, an accent identification experiment showed Spanish listeners have significantly more difficulties identifying L1 and L2 English accents, with the exception of their own accent, compared to listeners primarily from mid-Europe West-Germanic countries, such as Germany and the

Netherlands (Kristiansen, Zenner & Geeraerts, 2018). The authors claimed that this can be linked to an interplay of factors, such as perceived accent strength (the stronger the accent, the more easily recognizable), limited familiarity with the L1 of the L2 English speaker, but also with a lack of familiarity with L1 Englishes (due to, for example, ineffective foreign language education, see Caraker, 2016). Taken together, the greater linguistic distance that exists between Dutch and Spanish, the likely lack of familiarity of Spaniards with Dutch, as opposed to Dutch and German, and the differences between these countries in English fluency levels and exposure to L1 Englishes, might negatively affect the ability for L2 English speakers from Spain to identify Dutch-accented English and understand it.

Singapore, an important trading partner to the Netherlands as well. The Confederation of Netherlands Industry and Employers (VNO-NCW, 2018), presents an interesting contrast to the Netherlands, Germany, and Spain regarding the status of English. It is an island state populated primarily by three major ethnic groups with their own linguistic backgrounds (Chinese: 73%, Malays: 13.3%, and Indians: 9.1%) that have been using English as an L2 for centuries and are considered highly fluent in English (as a nation ranked number 3 out of 88 nations, EF, 2018; Departments of Statistics, Singapore, 2018). The dominant position that English as an L2 has acquired in Singapore, originally due to British colonialization, has led to the official recognition of English as a national language in Singapore, in addition to Malay, Mandarin English, and Tamil. English in Singapore has gone through a process of 'nativization', which means that the presence and use of English in Singaporean society has been intense enough to develop into a variety referred to as Singaporean English that is marked by a pronunciation, vocabulary, and grammar that is distinguishable from L1 and other L2 varieties of English. The emergence of varieties of English, such as Singaporean English, has been discussed in the context of the status of Englishes from most former British colonies in Asia, Africa, and the Caribbean. In general, such varieties, like Singaporean English, are not defined as L1 Englishes similar to, for example, British or American English because they did not originate in the nations that from a western and European perspective were considered the cultural and linguistic centers from which English emerged and developed during the British colonial period. In addition, they are not norm-providing in the sense that they are not used as a model for English learning by English learners outside of their own nation or even within their own countries (Kachru, 1982; Bolton & De Costa, 2018; Trudgill & Hannah, 2008). Indeed, while an increasing number of Singaporeans view English as their L1 (Departments of Statistics, Singapore, 2018; Lee, 2016), and English is an official language in education, academia, law, government, and business,

it has a separate, functional, status from Malay, Mandarin or Tamil. The latter are languages from different language families than English, that mark Singaporeans' cultural identity (Malay, Chinese or Indian) and are mostly used in non-formal, private contexts (Departments of Statistics, Singapore, 2018; Lee, 2016). This suggests that for Singaporeans, dealing with speakers of English with various cultural and linguistic identities is commonplace (Bokhorst-Heng & Caleon, 2009), which makes them part of the global LFE speech community as well as members of their own 'local' LFE speech community. As a result, Singaporeans might have developed a more flexible language norm that does not necessarily view English production that is not L1 as a potentially negative reflection on the speaker or a hindrance to speech understanding and effective information transfer. Therefore, studying how Singaporean speakers of English respond to L2 English accentedness (in our case Dutch-accented English) compared to L2 English speakers such as Germans and Spaniards can offer interesting insights into how different English speaker groups, and LFE community members, understand and evaluate (Dutch-accented) L2 English.

5.1.3 Speech understandability L2 English accents

If an LFE speech community, with its own language norms and attitudes, does indeed exist, it needs to be established how well L2 English varieties produced by different speaker groups in that community are understood and, consequently, how effective L2 English interactions are. There have been many studies that have researched the degree to which speech is (perceived to be) understood (e.g Hendriks et al., 2016, Munro & Derwing, 1995a,b; Derwing & Munro, 1997; Munro, Derwing & Morton, 2006; Hendriks et al., 2017; Hendriks et al. 2018; Gerritsen et al., 2010; Major et al., 2005; Nelson, 2011; Yorkston et al., 1996). Most studies have assessed understanding in terms of one distinct level of comprehension with research methods that allow for one component of speech understanding to be studied, for example, through orthographic transcription (e.g. Nelson, 2011; Yorkston et al., 1996) which reflects the degree to which listeners are able to determine individual words, or asking content questions to test whether content has been understood and to what extent (e.g. Major et al., 2005). These research methods suggest that understanding speech consists of distinct components.

In L2 English interactions, people are likely to have varying linguistic backgrounds and might not always be familiar with the particular English variety or English accent of the speakers they engage with. These circumstances might negatively affect understanding at different levels, that is, in terms of a person's ability to distinguish words and phrases, or their ability to understand the meaning (i.e. to comprehend)

or possibly the purpose (i.e. to interpret). These circumstances make it relevant to design L2 English research in such a manner that potential misunderstandings in terms of all these components of understanding can be investigated. In an attempt to measure speech understanding as a multi-component phenomenon, Nejjari et al. (Nejjari et al., 2012; Nejjari et al., to be published) operationalized Kachru and Smith's (Kachru & Smith, 2008) concept of speech understanding as a process consisting of three levels. The first level is *intelligibility*, which refers to the manner in which utterances are deciphered into individual sound patterns that form words and sentence-level elements. *Intelligibility*, as mentioned above, can be measured by asking listeners to orthographically transcribe individual words or sentences produced by speakers (see also Nejjari et al., 2012; Major et al., 2005; Nelson, 2011). The second level is *comprehensibility*, which is the manner in which words and sentences are understood when it comes to the individual meaning of words and how words put together express meaning within a specific context (see Nejjari et al., 2012; Nejjari et al., to be published). *Comprehensibility* can be measured by asking listeners content questions on speech samples that require that they understand the meaning of what was communicated. Kachru and Smith (2008) indicate that unlike *intelligibility*, *comprehensibility* requires that a listener understands the syntax, semantics and the physical context in which an utterance is heard (Wang 2007; Yorkston, Strand, & Kennedy, 1996). The third level is *interpretability*, which is difficult to distinguish from *comprehensibility*, because both deal with meaning beyond recognition of sound patterns that form words and phrases (Orikasa, 2016). *Interpretability* refers to whether listeners are able to grasp a speaker's intentions and the cultural baggage that is required for discourse strategies to be understood or correctly interpreted (see Nejjari et al., 2012; Nejjari et al., to be published). Nejjari et al. (2012) has shown that Dutch-accented English hindered *intelligibility* but not *comprehensibility* and *interpretability* for British listeners; however, when this experiment was replicated with Dutch listeners, the three components of *speech understandability* were not negatively affected by Dutch-accented English (Nejjari et al. to be published). The researchers suggest that this might have been a result of the familiarity Dutch listeners had with Dutch-accented English. In the context of the present study, it could be the case that listeners who are probably not familiar with Dutch-accented English, such as the Spanish and Singaporean listeners, might have more difficulties with understanding Dutch-accented English than German listeners who, due to their geographical proximity (and the linguistic relatedness of German and Dutch) are likely more familiar with Dutch-accented English. In order to further study the manner in which Dutch-accented English is understood by

Germans, Singaporeans, and Spaniards, Nejjari et al.'s (Nejjari et al., 2012; Nejjari et al., to be published) operationalization of Kachru and Smith's (2008) three components of *speech understandability* will also be included in the current experiment.

5.1.4 Speaker evaluations

In addition to studying speech understanding to gain more insight into L2 English interactions, it is also relevant to study the *speaker evaluations* L2 English accents arouse in terms of the characteristics of its speakers, such as intelligence, friendliness or assertiveness. Some studies have shown that *speaker evaluations* of L2 accented English are not necessarily more negative compared to *speaker evaluations* of L1 accented speech in terms of *affect* (e.g. friendliness, likeability) or *dynamism* (e.g. enthusiastic, proactive) (e.g. Hendriks et al., 2016; Hendriks et al. 2017; Hendriks et al., 2018; Nejjari et al., 2012; Nejjari et al., to be published). At the same time, several studies have shown that L2 (and L1) speakers of English tend to ascribe speakers of L2 English lower *status* (e.g. intelligent, cultured, competent) (e.g. Dalton-Puffer et al., 1997; He & Zhang, 2010; Ryan & Sebastian, 1980; Ryan & Bulik, 1982; Cargile 1998; Lindemann, 2002; McKenzie, 2008; Matsuura, Chiba & Yamamoto, 1994). These varying results might be a result of the L2 listeners' linguistic circumstances. For example, Singaporeans might have less negative *speaker evaluations* of Dutch-accented English compared to Germans and Spaniards since in their society English is an officially recognized language, used in all educational, professional, and political contexts as the primary tool to communicate across ethnic boundaries and therefore Singaporeans might be used to and more accepting of various English accents and varying English fluency levels. It is therefore important to study the *speaker evaluations* of L2 English accents by listener groups with particular linguistic circumstances to better understand their impact on the perceptions of speakers. Providing more insight into the influence of language use on listener perceptions beyond understandability, reflecting a more holistic view of SLA (see also above), would be seem particularly relevant in an LFE community.

5.1.5 Communication context

L2 English is not only spoken across and within L2 English speaker groups with diverse L1 backgrounds, it is also used in various different communication contexts. In different contexts, communication has diverse purposes and is linked to expectations in terms of communication behaviors. If a speaker violates these expectations, in a negative or positive sense, this can affect how they are perceived by listeners. This process of either negative or positive expectation violations was defined as

Expectancy Violation Theory (Burgoon, Stacks & Woodall, 1979; Burgoon & Burgoon, 2001) and applied to non-verbal communication behavior in, for example, intercultural personal communication contexts. If we apply this theory in an LFE speech community setting to accent production, it might explain differences in how evaluations of accents and speakers can vary across communication contexts. It might be the case that certain accents (L1 or L2) are deemed desirable or acceptable in specific communication contexts, but not in others. Cargile (1997) investigated American (Anglo- and Asian American) listeners' evaluations of Mandarin-accented English compared to standard American-accented English in a higher education lecture context compared to a job interview context. These communication contexts can both be considered high stakes contexts, since not understanding a speaker's accent in these contexts might mean that content is not transferred effectively; for example, in the case of a lecture, students might not be able to perform optimally on tests as a result (lecture). Another example of the high stakes nature of such contexts is that evaluating an interviewee negatively on the basis of their accent may result in that person not being hired for a position (job interview). Cargile (1997) found that listeners made no distinction between standard American-accented English and Mandarin-accented English in a job interview context, but did so in a higher education lecture context, where the Mandarin-accented English aroused lower dynamism, status, and affect than the standard American English accent. Cargile, whose study involved students as listeners, suggests that listeners' judgments might have been influenced by the fact that they were more familiar with the lecture context than the job interview context.

In another study that tested responses to accents in different communication contexts, Nejjari et al. (to be published) investigated Dutch listeners' responses to Dutch-accented English compared to standard British and American English in three communication contexts that these listeners would likely be familiar with in international settings (education, business, tourism). The results showed that communication context impacts speaker evaluations. The speakers evaluated in a job pitch context were almost universally perceived more negatively, regardless of their accent, an effect which was not observed for the lecture and the audio tour. The results differ from those of Cargile (1997), potentially because the listeners in Cargile (1997) were L1 English speakers, and in Nejjari et al. (to be published) they were L2 English speakers.

The influence of context emerging from these studies might be an indication that different language norms hold in different communication contexts, as suggested by Expectancy Violation Theory's negative or positive expectation violations (see above). This, in turn, illustrates the relevance of communication context as a factor

in accentedness studies; accents can evoke different judgments from listeners in differing contexts. In the present study, therefore, communication context will be included as a variable to compare responses by L2 English speakers to L2 and L1 English accents in different situations of use.

5.1.6 Purpose experiment

The current study investigated the reactions to L1 and L2 accented English by listeners from Germany, Spain, and Singapore. It featured Dutch-accented English and standard British and American English accents respectively, as the latter are the L1 Englishes that are globally the main models used in English education for L2 English speakers (British Library, 2019), and therefore function, from a traditional SLA perspective, as the 'target' language that learners are encouraged to master in formal education. If as many learners as possible acquire this 'target' language, this will increase the likelihood of speakers with different L1 backgrounds understanding each other, even though some researchers believe that non-nativeness does not have to hinder understanding (e.g. Kachru, 1982; Alexander, 1999; van Oostendorp, 2002; Jenkins, 2006; Seidlhofer, Breiteneder & Pitzl, 2006). However, it remains unclear what the actual effects of L2 English accents are within the LFE speech community on speech understandability and speaker evaluations, and in different communication contexts, which is a question the present study aimed to answer with respect to Dutch-English. If L2 English speakers in different countries do not necessarily share the traditional SLA perspective, and L2 English accents do not hinder speech understandability and speaker evaluations, this might mean that in terms of accentedness, achieving an L1 English accent is not necessary.

The present study investigated the response to Dutch-English (versus British and American) accent by L2 English groups from countries with varying English fluency levels that are important partners to the Netherlands, namely Germany, Spain, and Singapore. Furthermore, the groups have L1s that vary in relatedness to Dutch and/or English. While German, Dutch, and English are linguistically related to one another because they are West Germanic languages, Spanish is not as related to Dutch as German or English, since it is a Romance language. Singaporean listeners, in turn, speak a variety of English that is generally not considered an L1 English, but is heavily influenced by British English (Brown, 1999). This might mean that Singaporean English is linguistically more related to English, Dutch, and German compared to Spanish. In the present study, the different characteristics of the three listener groups were expected to lead to different responses to Dutch-accented English. For example, studies have shown that familiarity can aid *speech understandability* (Nejjari et al.,

2012; Major et al., 2005; Fayer & Krasinski, 1987; Bent & Bradlow, 2003; Varonis & Gass, 1982; Wang, 2007; Wang & van Heuven, 2007). As German listeners are likely to be more familiar with Dutch-accented English, because the Netherlands is a neighboring country to Germany, they are one of the closest political, economic and EU partners (Federal Foreign Office, Germany, 2019), and both have a West Germanic language as their official language and L1, they might, as a result, have higher *speech understandability* of Dutch-accented English compared to Spaniards or Singaporeans who are likely to be less familiar with Dutch and a Dutch English accent.

The varying results from earlier studies with regard to the *speech understandability* and *speaker evaluations* of L2 English, and the limited research into the manner in which communication context might impact responses to L2 English (e.g. Nejjari et al., to be published; Cargile, 1997), led to the inclusion of three different communication contexts in our experiment: a lecture, a retail manager job pitch, and an art gallery audio tour. These contexts were selected to represent the settings in which L2 English is frequently employed: academia, international business, and tourism (Nejjari et al., to be published; Cargile, 1997; Nickerson, 2005).

In order to determine whether accent and context affect *speech understandability* and *speaker evaluations*, two research questions and six expectations were formulated. The first research question was:

RQ1: Do German, Spanish, and Singaporean listeners (*listener group*) display different *speech understandability (intelligibility, comprehensibility, interpretability)* in response to Dutch-accented English compared to standard British and American English accents (*accent*), and does *context* (lecture; audio tour; job pitch) affect their responses?

Due to Spain's reported lower English fluency level, and the fact that Spanish is typologically more distant from Dutch compared to German, it is expected that compared to German listeners, Spanish listeners will show lower *speech understandability* of Dutch-accented English, standard British and American English accents compared to German and Singaporean listeners. Therefore, the following expectation was formulated:

Expectation 1a: Spanish listeners will display a lower level of speech understandability of the tested accents compared to German and Singaporean listeners.

Singapore's reported high English fluency, its official recognition of English as a national language, and its linguistically diverse population that regularly communi-

cates across ethnic groups in English might facilitate *speech understandability* of the various English accents in the Singaporean listener group more than in the listener groups from the relatively homogenous societies in this study, namely Germany and Spain. Consequently, the following expectation was formulated:

Expectation 1b: Singaporean listeners will display a higher level of speech understandability compared to German and Spanish listeners.

Following Cargile (1997) and Nejjari et al. (to be published) our experiment assessed responses to accents in a lecture, job pitch, and an audio tour context. Only highly educated listeners were selected to assess the accents, since they represent the population most likely to be relatively familiar with the three contexts. The communication context a highly educated listener was assumed to be most familiar with was the lecture context, because individuals who followed high(er) education are very likely to have attended academic lectures regularly. This expected stronger familiarity with the lecture context was expected to aid *speech understandability* in this context, because listeners were expected to recognize the lecture genre and the communicative rituals that the genre entails, which could allow them to more easily concentrate on the content of what is communicated, even if that content is not familiar to them. Therefore we expected that:

Expectation 1c: The lecture communication context will evoke higher speech understandability compared to the job pitch and audio tour communication contexts.

In addition to studying the effects of accent and context on *speech understandability*, *speaker evaluations*, or listeners' evaluations of a speaker's characteristics, were assessed as well. It was assumed that the listener groups' responses might differ due to their diverse linguistic backgrounds. This, in turn, could be an indication of potentially different language norms. As a result, the second research question was:

RQ2: Do German, Spanish, and Singaporean listeners (*listener group*) display different *speaker evaluations (status, affect, dynamism)* in response to Dutch-accented English compared to standard British and American English *accents* (accent), and does *context* (lecture; audio tour; job pitch) affect their responses?

In general, speaker evaluation research has shown that L2 listeners ascribe L2 English accents lower status, but not lower affect and dynamism, compared to L1

English accents. This is why we expected that L2 English speakers from Germany and Spain would ascribe Dutch-accented English lower status than standard British and American English:

Expectation 2a: German and Spanish listeners will ascribe Dutch-accented English lower status compared to standard British and American English accents.

Singapore's linguistically and culturally diverse population has developed its own nationally recognized variety of English, Singaporean English, which is generally not considered an L1 English, since it is not norm-providing to learners of English outside of Singapore. Furthermore, it was assumed that Singaporeans communicate with English speakers from various national and linguistic backgrounds, and that therefore, Singaporean listeners might not view accentedness as an important marker of a speaker's character or abilities. As a result, they might not evaluate L2 and L1 English accents differently. This is why we expected that:

Expectation 2b: Singaporean listeners will not display different speaker evaluations in response to Dutch-accented English compared to standard British and American English accents.

Based on the results of a limited number of studies, it would seem that L1 and L2 listeners evaluate speakers with the same accent differently when the accent is presented in different communication contexts (*context*). In the case of L1 English speakers as listeners (Anglo- and Asian Americans) [12], a lecture aroused lower dynamism, status, and affect when produced with an L2 English accent compared to an L1 English accent. In a job interview context this effect was not observed. In a study by Nejjari et al. [8], involving the same contexts as the present study, and L2 English speakers as listeners (Dutch), no such effect was found for the lecture context and L1 and L2 English accents. However, the job pitch, regardless of accent, evoked significantly lower *speaker evaluations* compared to the lecture and audio tour. As the current study is also focused on L2 English listeners, albeit on three different groups of L2 English speakers (German, Spanish, Singaporean), we expect similar speaker evaluation patterns to emerge as in Nejjari et al. (to be published), resulting in the final expectation:

Expectation 2c: The job pitch context will evoke lower speaker evaluations compared to the lecture and audio tour context.

5.2 METHOD

To investigate the effects of *listener group*, accent and *context* on *speech understandability* and *speaker evaluations*, a matched-guise experiment was conducted in which we compared three listener groups' responses (N=1699) from Germany (N=617), Spain (N=540), Singapore (N=542) to three English accents (Dutch-accented English, standard British and American English) in three communication contexts (a lecture, an audio tour, a job pitch). All listeners responded to stimuli (speech samples) via an online questionnaire. The experiment had a within-subject multifactorial design. All listeners (*listener groups*) were exposed to the independent variables (*accent*, *context*) and evaluated the stimuli on the dependent variables (*speech understandability, speaker evaluations*).

5.2.1 Speakers: matched-guise speaker, control and filler speakers

To avoid responses to the voice characteristics of individual speakers, a matched-guise speaker was selected to produce the three accents. The matched-guise speaker had been assessed in an earlier speech evaluation experiment (Nejjari et al., 2019), which showed that he could produce the three accents under study that represented the independent variable accent: (1) standard British English, (2) standard American English, and (3) the typical English accent of highly educated L1 speakers of Dutch. In the current study, the standard accents of British and American English refer to accents generally associated with the national accent norm of these nations; Dutch English does not have an explicit national norm. We regard a typical Dutch English accent in the present study as containing features that L1 speakers of Dutch and others familiar with Dutch and Dutch English will recognize as such. For example, because Dutch lacks dental consonants [ð] as in *this, mother, breathe* and [θ] as in *think, Martha, breath*, they are often mispronounced as stop consonants, [d] and [t] respectively, by Dutch speakers of English. Dutch also lacks voiced fricatives and plosives in the coda, causing the voiced obstruents of English to generally be pronounced as their voiceless counterparts in Dutch speakers' English (e.g. *live, badge, bad, bag* will be said with [f, tʃ, t, k] (Gussenhoven & Broeders, 1997). As no standard has been defined for a standard Dutch English accent, in this study it is defined as being 'typical' (Nejjari et al., 2019, to be published).

To prevent listeners from deducing that they were listening to the same matched-guise speaker a number of times, we included speech samples from six male control speakers as stimuli: two L1 speakers of standard British English, two L1 speakers of standard American English, two L1 speakers of Dutch who have a typical

Dutch accent in English. All but one had been assessed on the representativeness of their accents in our previous study (see Nejjari et al., 2019). The speaker who was not assessed was one of the Dutch-accented English control speakers, who was regarded by experienced linguists as a representative speaker of Dutch-accented English. One further speaker produced a speech sample in a standard British English accent that was presented to listeners at the beginning of the experiment (the filler speech sample) to familiarize them with the task. The filler speaker had been assessed in our earlier study by L1 speakers of British English as an L1 and standard speaker of British English (Nejjari et al., 2019). All speakers were aged 35 to 60 at the time of recording, had at least a master's degree, and were English language and/or linguistics specialists in some capacity.

5.2.2 Stimuli

One filler text (on a general topic) and three texts that represented the independent variable *context* were used as the basis for the speech samples: (1) an introduction to a marketing lecture; (2) an art gallery audio tour segment; (3) a job pitch for a retail management position (see section 5.5.1). All but the filler text reflect three contexts in which LFE is commonly used: higher education, tourism, and international business (Edwards, 2016; Nickerson, 2005; Gerritsen et al., 2016). The lecture, audio tour, and filler texts were selected from an IELTS Academic English test and the job pitch text from a human resources webpage. The matched-guise speaker produced the three accents in all three contexts, resulting in nine speech samples. The six control speakers produced their L1 accent in the three contexts, resulting in 18 samples. The filler speaker produced one speech sample on a general topic in standard British English.

5.2.3 Listeners: age, education, L1 language(s), English fluency

Table 1 shows the listener groups' sex, average age, English fluency, and education level.

Highly educated listeners with at least average self-reported English fluency were selected to represent the listeners who would most likely communicate in English in the selected communication contexts with other L2 speakers or L1 speakers of English. For example, in academia, reflected in an increasing number of English-taught academic degree programmes. In addition, in international business, which requires that people in the Netherlands, Germany, Spain and Singapore increasingly need to perform in English with international business contacts (Nickerson, 2005; Bouma, 2016; Bronkhorst, 2015; van Gaal, 2018).

Table 1. Listeners (N=1699): age, % sex, self-reported English fluency, education level.

	Germany (N=617)	Spain (N=540)	Singapore (N=542)
Mean age	38 (Min=19; Max=83)	37 (Min=18; Max=64)	34 (Min=18; Max=80)
Male	46.0%	34.3%	44.9%
Female	54.0%	56.7%	55.1%
Mean self-reported English fluency (Min=3; Max=5)[a]	4.03	3.62	4.24
Education level[b]			
Undergraduate	0.0%	0.0%	27.5%
Bachelor	43.3%	55.9%	85.2%
Master	52.0%	39.4%	12.8%
Doctorate	4.7%	4.1%	2.0%

[a]Mean self-reported English fluency was the mean for indicated levels for English listening, reading, writing, speaking skills on a 5-point scale. The mean had to be at least 3 in order to participate in the questionnaire (1: very low; 2: low; 3: average; 4: high; 5: like a L1 speaker).

[b]The original categories were: education at A-level, undergraduate; bachelor, master, PhD, other. Only listeners who had reached at least an undergraduate degree level of education were allowed to participate in the questionnaire.

5.2.4 Instrumentation

Each listener evaluated four different speech samples by four different speakers: the filler sample followed by three experimental samples (in this order: Lecture, Audio Tour, Job Pitch) produced by the matched-guise speaker and the control L1 speakers. To ensure that the nine (plus 1 filler) matched-guise samples could be evaluated in each context for each accent, to avoid repeating the content of each context, and to limit any order effect, 18 listener groups were created. Data was collected from 18 listener groups (consisting of at least 30 listeners each) in each country. This procedure and subsequent data quality checks (see also section 5.2.4) resulted in 617 completed questionnaires for Germany, 540 for Spain, and 542 for Singapore.

Speaker evaluations and *speech understandability* were assessed in the questionnaire followed by questions regarding the listeners' English language skills, L1 languages, and general personal details. First, the listeners were provided with a general introduction to the questionnaire and asked to click on a link to listen to a speech sample. They were then asked to answer the *speaker evaluations* questions, followed by the *interpretability* and *comprehensibility* questions to assess the two

dimensions of *speech understandability*. Subsequently, the listeners were asked to click on another link which led them to the first 10-12 words of the relevant speech sample, which they could listen to a maximum of two times, in order to assess *intelligibility* by means of an orthographic transcription task. The questionnaire was designed in this manner to ensure that listeners would provide their first impression of the speakers' traits (*speaker evaluations*) and display their understanding of the speech samples' content. The *intelligibility* task required that the listeners pay close attention to what was stated in the selected segment of each speech sample, and therefore was assessed as a separate task and positioned after the *speaker evaluations* and *interpretability* and *comprehensibility* listening tasks.

To establish whether listeners could distinguish between the two L1 and the one L2 English accents, they were asked in an open question to indicate what they believed the speaker's country of origin was (see section 5.5.2 for their answers). This information was used to calculate whether the listeners' answers might have impacted their responses and to understand whether correctly or incorrectly identifying the speaker's accent affected *speaker evaluations*. For German listeners, 40.1% correctly indicated that the Dutch English speaker was from the Netherlands; for the Spanish and Singaporean listeners this was less than 2%. For each listener group, at least 87% recognized the standard British and American accent as L1 English accents. Dutch-accented English was recognized as an L2 English accent (i.e. not as specifically from the Netherlands) by a majority of the Germans (74%) and Singaporeans (69%). Only 42% of the Spanish listeners indicated that the Dutch-accented English speech samples came from a speaker who was an L2 English speaker.

5.2.5 Speech understandability

Following Smith and Nelson (2006), Kachru and Smith (2008), Nejjari et al. (2012, to be published), Bayyurt (2018), and Berns (2018) three questions were used to measure *speech understandability* and more specifically the ability (1) to literally recognize words (*intelligibility*), (2) to understand the meaning of the words within the context (*comprehensibility*), and (3) to understand the intention of the speaker/purpose of the message (*interpretability*).

To measure *intelligibility*, listeners were presented with a speech sample consisting of the first 11-12 words of the three sample texts and asked to write down literally what was stated. If respondents were able to do so, this was counted as intelligible. An Intraclass Correlation Coefficient was calculated to measure inter-rater reliability and showed a correlation for each listener group of .91 (Germany), .97 (Spain), and .95 (Singapore) (0.00 - 1.00), which is extremely high. To measure

comprehensibility, listeners were asked to indicate whether a statement on the content of each sample was correct or not. *Interpretability* was measured by having listeners indicate whether a statement on the communicative intentions of the speaker of each sample was correct or not.

5.2.6 Speaker evaluations

To assess *speaker evaluations*, listeners indicated on 5-point Likert scales (1=strongly disagree; 5=strongly agree; 3=neither disagree nor agree) to which extent they believed the speaker possessed 11 personality traits. These traits represent three dimensions of *speaker evaluation*, namely *status* (competent, educated, having authority, intelligent and cultured), *affect* (considerate, pleasant and friendly), and *dynamism* (energetic, enthusiastic, confident). The traits associated with *status* and *affect* are based on Nejjari et al. (2012). *Status* represented the degree to which a speaker was viewed as being intelligent and well-educated, and *affect* represented the degree to which a speaker was perceived as being likeable. *Dynamism* measured the impression the speaker gave of himself, also known as a speaker's 'self-presentation', and was based on Grondelaers, van Gent et al. (2015).

5.2.7 Data collection procedures

The experiment was conducted in 2017 and 2018 via Qualtrics, a global survey software and online data collection company that caters for (non-)commercial organizations, and was originally launched for academics and the complex requirements of research (Qualtrics, 2019). In our case, Qualtrics was requested to sample listeners who were L1 speakers of the main national language of Germany (German), Spain (Spanish), and Singapore (Singapore English) and were highly educated (i.e. had at least reached or completed undergraduate level education). No restrictions were placed in terms of regions in the three countries where the experiment was circulated. At least 540 listeners per country were needed, and participants were registered individually. The agreement with the service also allowed for the main researcher to check the data and request replacement listeners in the case of inadequate responses.

Of the German listener data, 5% was initially collected in the context of a cross cultural communication research course at Radboud University in the Netherlands. Qualtrics was hired to collect the remaining 95% of the data. Subsequently, the listener data from Spain and Singapore was also collected by Qualtrics. On average, listeners needed 15 minutes to complete the questionnaire. For each country the collected data was checked by the first author, which resulted in three data collection rounds that took about two weeks, with approximately 35% of the total data were excluded.

The reasons for excluding data included: providing nonsense answers or numbers, symbols, or only providing neutral (mid-scale) answers to all scale questions. All excluded data was replaced by new data which were checked again and replaced if necessary. Response analyses were conducted on the data collected for the matched-guise speech samples only.

5.2.8 Statistics

Descriptives and frequencies were calculated to establish means and percentages of listener characteristics and responses. Varimax rotations, analyses of variance, Pearson correlations, logistic regressions were calculated to establish the interactions, effects and correlations of the defined variables.

5.3 RESULTS

5.3.1 Speech understandability

This section will focus on whether German, Spanish, and Singaporean listeners display different *speech understandability* in response to Dutch-accented English compared to standard British and American English accents, and if context affects their responses (RQ1). With regard to our expectations, we will report whether the Spanish listeners displayed a lower level of *speech understandability* of the tested accents compared to German and Singaporean listeners (expectation 1a); whether Singaporean listeners displayed a higher level of *speech understandability* compared to German and Spanish listeners (expectation 1b); and whether the lecture communication context evoked higher *speech understandability* compared to the job pitch and audio tour communication contexts (expectation 1c). The frequencies and means for *speech understandability* (three accents, in three contexts) are presented in sections 5.5.3, 5.5.4 and 5.5.5. A summary of the significant main and interaction effects are provided in Table 2.

5.3.2 Intelligibility

The listeners' transcriptions of the first 11 to 12 words of each speech sample could result in a maximum of 11 (Lecture, Job Pitch) or 12 (Audio Tour) correctly transcribed words. Scores were equalized by computing the proportions of correct words. On average, the German listeners (N=617) were able to correctly transcribe 7.47 (SD=3.82) words, the Singaporean listeners (N=542) 6.84 (SD=4.08) words, and the Spanish listeners (N=540) 6.14 (SD=3.17) words. The results for *intelligibility* in relation to accent and context are shown in Figures 1 and 2, with 95% confidence intervals.

Figure 1. Mean scores of intelligibility for accent (British English, American English, Dutch English) and listener group (Germany, Singapore, Spain).

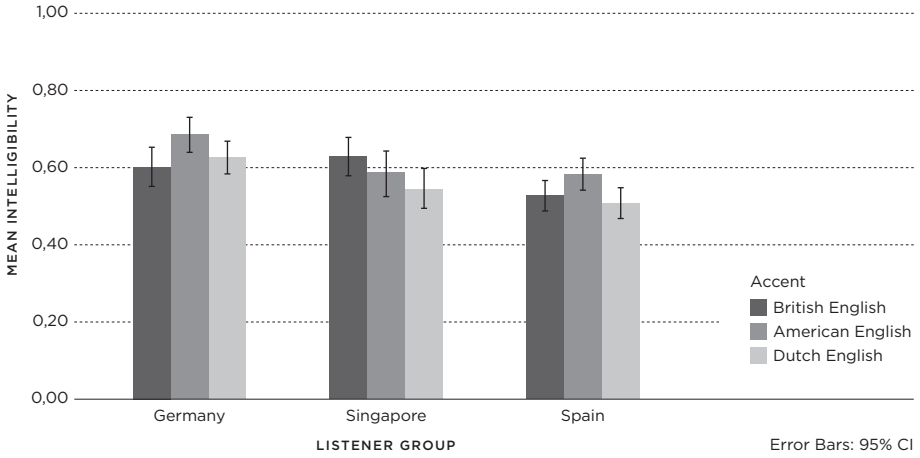
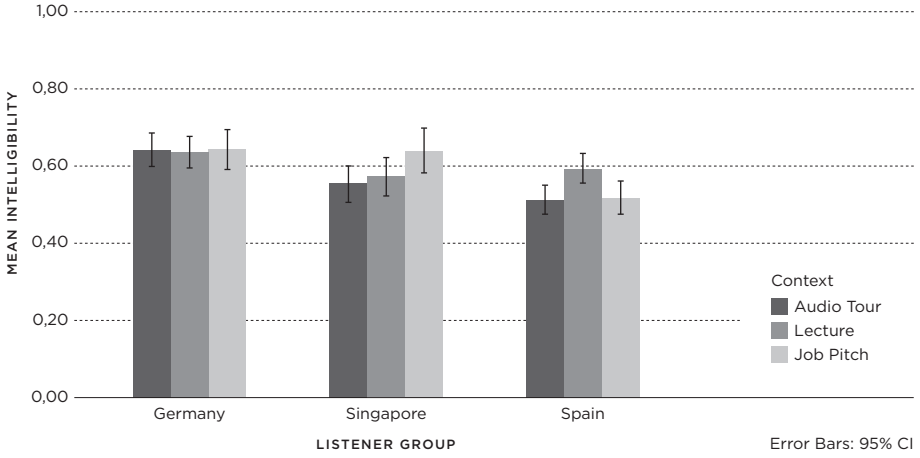


Figure 2. Mean scores of intelligibility for context (Audio Tour, Lecture, Job Pitch) and listener group (Germany, Singapore, Spain).



In terms of *accent*, there was substantial overlap between the confidence intervals of the bars, indicating that the responses to the three accents did not differ greatly between the three listener groups (Figure 1). In terms of *context*, there were substantial differences in the confidence intervals between and within the three *listener groups*, indicating that the responses to the three contexts differed more strongly (Figure 2).

A univariate analysis of variance was applied to investigate the effects of *listener group*, *accent*, and *context* on *intelligibility* and their interaction effects. The three-way interaction was not significant ($F(8, 1672)=1.449$, $p=.171$, $PES=.007$), and neither was the two-way interaction between *accent* and *context* ($F(4, 1672)=0.167$, $p=.995$, $PES=.000$). The two other interaction effects were significant though not very outspoken: *accent* by *listener group* ($F(4, 1672)=2.539$, $p=.038$, $PES=.006$) and *context* by *listener group* ($F(4, 1672)=2.450$, $p=.044$, $PES=.006$). The interaction effects can be interpreted by analysing the main effects, one of which, *context*, was not significant ($F(2, 1672)=2.344$, $p=.096$, $PES=.003$). Accent ($F(2,1672)=4.704$, $p=.009$, $PES=.006$) and *listener group* were significant ($F(2, 1672)=15.932$, $p=.000$, $PES=.019$).

More details about the main effects were obtained through separate analyses of variance for each *listener group*, the factor involved in both significant two-way interactions. In the German group, effects were not significant for *context* ($F(2, 608)=.148$, $p=.862$, $PES=.000$), and *accent* by *context* ($F(4, 608)=.315$, $p=.263$, $PES=.009$). However, the effect of *accent* was significant ($F(2, 608)=3.679$, $p=.026$, $PES= .012$), with a post-hoc difference (Tukey's HSD) between *intelligibility* of American English (the highest score, .686) and British English (the lowest score, .602). In the Singaporean group, no significant effects were found (*accent*, $F(2, 533)=2.555$, $p=.079$, $PES=.009$; *context*, $F(2, 533)=2.793$, $p=.062$, $PES=.010$; *accent* by *context*, $F(4, 533)=.342$, $p=.849$, $PES=.003$). In the Spanish group, the two main effects and one interaction were significant (*accent*, $F(2, 531)=3.860$, $p=.022$, $PES=.014$; *context*, $F(2, 531)=5.119$, $p=.006$, $PES=.019$; *accent* by *context*, $F(4, 531)=1.901$, $p=.109$, $PES=.014$). The post-hoc results for *accent* showed significantly higher *intelligibility* for American English (.584) than for Dutch English (.509). The post-hoc results (Tukey's HSD) for *context* showed a significant difference in *intelligibility* between Lecture (.592) and the lower scores of Audio Tour (.511) and Job Pitch (.517).

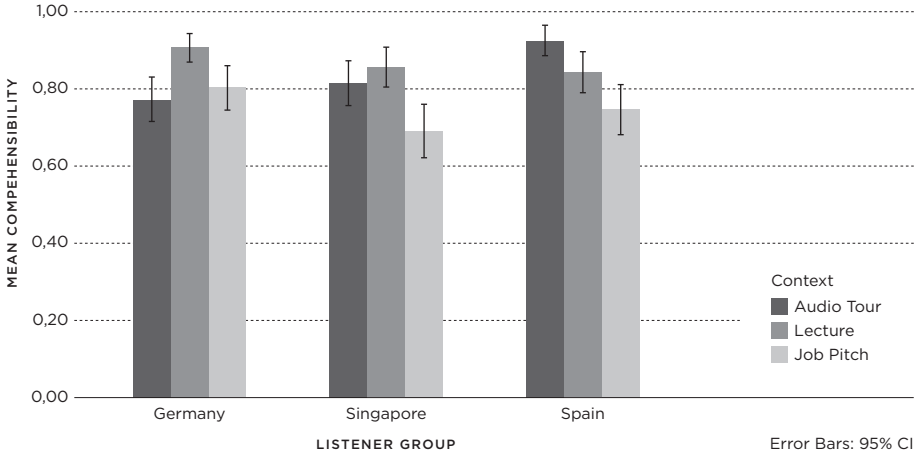
The results show that there is no consistent hierarchy of accents and contexts for *intelligibility*; instead, they vary within the *listener groups*. No general effect was found for accent; however, when higher *intelligibility* was observed, it was for American English. Furthermore, Dutch-accented English never yielded the lowest intelligibility. Given the modest effects of *accent* and *context* on *intelligibility*, we

conducted a post-hoc analysis (Tukey's HSD) on the main effect of *listener group*, which yielded significant differences between the groups: Germans (.634) > Singaporeans (.588) > Spanish (.540), which showed the German listeners achieved highest scores for intelligibility, followed by the Singaporean and Spanish.

5.3.3 Comprehensibility

Of the Spanish listeners (N=540), 84% correctly comprehended the speech samples, which for German listeners (N=617) was 83% and for Singaporean listeners (N=542), 79%. In terms of accent, the results demonstrated substantial overlap in the confidence intervals, indicating that there are no strong differences between the responses to the three accents per *listener group*, which is why no figure has been incorporated for *accent*. The results for *comprehensibility* in relation to *context* are shown in Figure 3, with 95% confidence intervals. In terms of context, there were more substantial differences in the confidence intervals between bars, indicating that the *comprehensibility* of the three contexts differed more strongly within each individual *listener group*.

Figure 3. Mean proportions of correct comprehensibility for context (Audio Tour, Lecture, Job Pitch) and listener group (Germany, Singapore, Spain).



A logistic regression was applied to test the effects of *listener group* (Germany, Spain, Singapore), *accent* (standard British English, standard American English, Dutch English), and *context* (Audio Tour, Lecture, Job Pitch), as well as their interaction. The final model contained significant effects of *listener group* (Wald=6.063, df=2, p=.048), *context* (Wald=19.476, df=2, p=.000) and their interaction (Wald=22.997, df=4, p=.000), but no significant effect of accent.

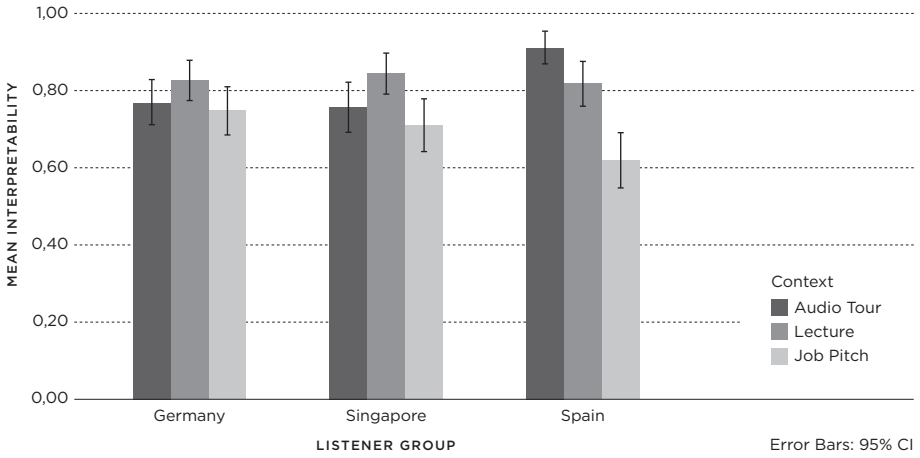
To investigate these effects in detail, we conducted a logistic regression per *listener group* to test the effect of context within each group. The German data returned a significant *context* effect (Wald=14.643, df=2, p=.001), with Lecture leading to significantly higher *comprehensibility* than Audio Tour and Job Pitch. The Singaporean data also returned a significant *context* effect (Wald=15.701, df=2, p=.000), but here it was the Job Pitch context that led to lower *comprehensibility* than the Audio Tour and Lecture contexts (see Figure 4). A significant context effect was also found for the Spanish group (Wald = 19.746, df=2, p=.000), with *comprehensibility* being significantly higher in the Audio Tour context than in the Job Pitch context. All pairwise comparisons for the Spanish listeners turned out to be significant, which gave a decreasing *comprehensibility* effect going from Audio Tour to Lecture to Job Pitch (see Figure 3), yielding the same pattern that was found for *interpretability*. In sum, the three contexts show varying results with respect to *comprehensibility* within the three *listener groups*. However, these are not very outspoken, which means that, in general, there is no one context that is much more difficult or easier to comprehend than another within the three *listener groups*.

Next, we applied a logistic regression separately to the three contexts with *listener group* as predictor. For Audio Tour, there was a significant *listener group* effect (Wald=15.884, df=2, p=.000), with higher scores for the Spanish compared to Germans (see Figure 2). For Lecture, there was no significant difference between *listener groups* (Wald=4.314, df=2, p=.116). Job Pitch yielded a significant *listener group* effect (Wald=6.063, df=2, p=.048), but here *comprehensibility* was lower for the Spanish than for the Germans. In sum, this means that there is no general *listener group* effect on *comprehensibility*, and that differences between the *listener groups* do not depend on accent, but on context.

5.3.4 Interpretability

The German (N=617) and Spanish listeners (N=540) answered the *interpretability* question correctly in 78.3% of all cases, and the Singaporeans did so in 77.1% of cases (N=542). In terms of accent, the results demonstrated substantial overlap in the confidence intervals, indicating that there are no strong differences between the

Figure 4. Mean proportions of correct interpretability for context (Audio Tour, Lecture, Job Pitch) and listener group (Germany, Singapore, Spain).



responses to the three accents within each individual *listener group*, which is why no figure has been incorporated for accent. The results for *interpretability* in relation to *context* are shown in Figure 4, with 95% confidence intervals. In terms of *context*, there were more substantial differences in the confidence intervals between bars, indicating that the *interpretability* of the three contexts differed more strongly within each individual *listener group*.

A logistic regression was applied to test the effects of *listener group* (Germany, Spain, Singapore), *accent* (standard British English, standard American English, Dutch English), and *context* (Audio Tour, Lecture, Job Pitch), as well as their interaction. The final model selected only contained the effects of *listener group* (Wald=7.292, df=2, p=.026), context (Wald=41.873, df=2, p=.000) and their interaction (Wald=23.715, df=4, p=.000). There was no significant effect of *accent*.

To investigate these results in detail, we carried out a logistic regression per *listener group* to test the effect of context within each group. No significant *context* effect was found for the German data (Wald=4.045, df=2, p=.132). However, a significant *context* effect was found for the Singaporean data (Wald=9.169, df=2, p=.010), which turned out to be a single pairwise difference between Lecture and Job Pitch,

with a higher score for Lecture (see Figure 4). The Spanish data also yielded a significant *context* effect (Wald=41.873, df=4, p=.000), with lower interpretability for the Job Pitch compared with the other two contexts (Lecture, Audio Tour). All pairwise comparisons for the Spanish listeners turned out to be significant, with a decreasing interpretability effect going from Audio Tour, Lecture, to Job Pitch (see Figure 4). In other words, the three contexts show varying results within the three *listener groups*, with none of the speech samples being generally more difficult or easier to interpret than the other ones across listener groups.

To evaluate the differences between the *listener groups* with regard to *interpretability*, we applied a logistic regression for each of the three contexts with *listener group* as predictor. For Audio Tour, there was a significant listener group effect (Wald=16.175, df=2, p=.000), with a significant difference between higher scores for Spain and lower scores for Singapore and Germany. For Lecture, no significant difference was found between *listener groups* (Wald=.433, df=2, p=.805). Job Pitch yielded a significant l*istener group* effect (Wald=7.293, df=2, p=.026), but here Spain was found to have the lowest score, which was significantly different from Germany, but not significantly different from Singapore. Overall, this means that there was no general effect of *listener group* nor of *accent* (Germany, Singapore, Spain) on *interpretability*, but that the differences between the *listener groups* depended on the *context* (Audio Tour, Lecture, Job Pitch) investigated.

5.3.5 Summary table speech understandability

Table 2 provides an overview of the results for *speech understandability* for the three tested accents and contexts for all three listener groups.

5.3.6 Speaker evaluations

This section will report on the question whether German, Spanish, and Singaporean listeners display different *speaker evaluations* in response to Dutch-accented English compared to standard British and American English accents, and whether context affects their responses (RQ2). Furthermore, with regard to our three expectations, we will report whether German and Spanish listeners ascribed Dutch-accented English lower status compared to standard British and American English accents (2a); whether Singaporean listeners did not display different *speaker evaluations* in response to Dutch-accented English compared to standard British and American English accents (2b), and whether the job pitch context evoked lower *speaker evaluations* compared to the lecture and audio tour context (2c). A summary of the significant interaction effects, main effects, and post-hoc tests are provided in Table 4 at the end of this sec-

Table 2. Summary results Speech Understandability (Intelligibility, Comprehensibility, Interpretability).

N=1699	Intelligibility	Comprehensibility	Interpretability
Listener Group (L)	Germany > Singapore > Spain	NS	NS
Accent (A)	NS	NS	NS
Context (C)	NS	NS	NS
L x A	Germany: AE > BE Spain: AE > DE	NS	NS
L x C	Spain: LE > AT, JP	Germany: LE > AT, JP Singapore: LE, AT > JP Spain: AT > LE > JP Spain AT > Germany AT Germany JP > Spain JP	Singapore: LE > JP Spain: AT, LE > JP Spain: AT > LE > JP Spain AT > Germany, Singapore AT Germany JP > Spain JP
A x C	NS	NS	NS
L x A x C	NS	NS	NS

A = Accent; AE = American English; AT = Audio Tour; BE = British English; C = Context; DE = Dutch English; JP = Job Pitch; L = Listener group; LE = Lecture; NS = Not Significant.

tion. The mean and frequency measurements of the *speaker evaluations* of the three accents produced in three contexts are presented in sections 5.5.3, 5.5.4 and 5.5.5.

A factor analysis, using a principle axis factoring extraction method with an Eigenvalue >1 criterion for factor extraction, followed by a varimax rotation, on the ratings of the personality traits that were measured to represent the three speaker evaluation constructs, showed a resolution into three factors for the German and Spanish listener groups, and two factors for the Singaporean *listener group* (see Table 3).

Table 3. Rotated Factor Matrix: factor loadings scores on 11 scales with three factors for listener groups (Germany, Spain, Singapore). Only loadings >.550 are reported.

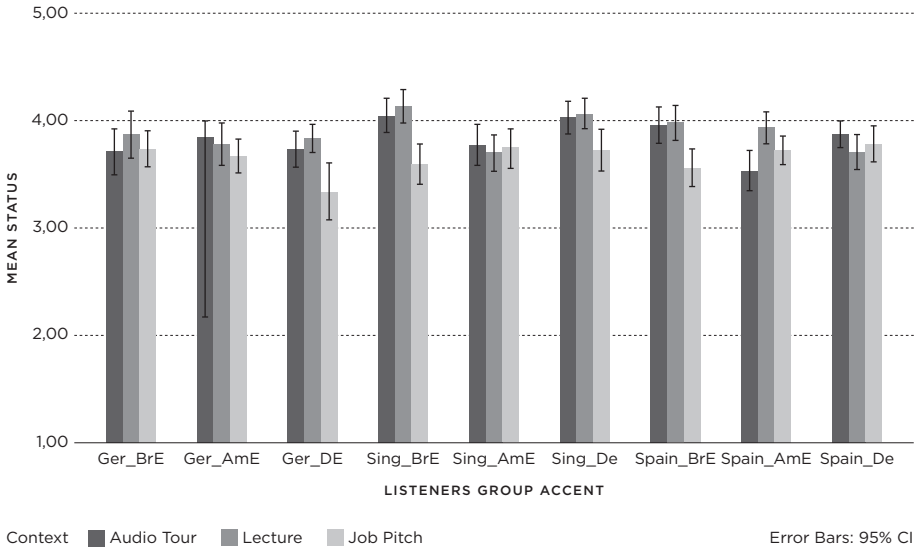
	Germany			Spain			Singapore	
Factor	1	2	3	1	2	3	1	2
Competent	.736			.770			.732	
Considerate					.667			.609
Cultured	.759			.826			.713	
Educated	.803			.741			.828	
Pleasant		.693			.774			.668
Energetic			.830			.846		.812
Authoritative			.780			.749		
Friendly		.844			.829			.813
Enthusiastic		.641				.675		.815
Intelligent	.779			.756			.763	
Confident	.683			.670			.717	

Status was analyzed for the personality traits competent, cultured, educated, intelligent, and *confident*. The personality trait *authoritative* was excluded, because it was not perceived as similar enough to the other *status* personality traits by any of the three *listener groups* and did not yield loadings above .550 in the Singaporean *listener group*. *Affect* was analyzed for the personality traits *pleasant* and *friendly*, because all three l*istener groups* evaluated these personality traits are part of one factor. The personality trait *considerate* was excluded because for German listeners it did not yield loadings above .550. *Dynamism* was analyzed for the personality trait energetic only, because for *listener groups* Germany and Spain this trait loaded as part of a third separate factor (see Table 3). This was not the case for Singaporean listeners; however, it was nonetheless analyzed as part of *dynamism*. The personality trait confident was excluded, because all *listener groups* regarded it as part of *status*. The personality trait *enthusiastic* was excluded because German and Singaporean listeners regarded it as part of *affect*, and Spanish listeners as part of *dynamism*. In sum, on the basis of the loadings (see Table 3), three factors were defined and used in further analyses: *status* (competent, cultured, educated, intelligent, confident), *affect* (pleasant, friendly), *dynamism* (energetic).

5.3.7 Status

The results for *status* in relation to *listener group*, accent and context are shown in Figure 5, with 95% confidence intervals.

Figure 5. Status (1=negative; 3=neutral; 5=positive) for Listener groups (Germany (Ger), Spain, Singapore (Sing)) per accent (British English (BrE), American English (AmE), Dutch English (DE)) and context (Audio Tour, Lecture, Job Pitch).



Univariate analyses of variance were applied and showed a significant three-way interaction for *listener group*, *accent* and *context* (F(8, 1672)=4.30, p=.00, PES=.02) and significant two-way interactions for accent and *listener group* (F(4, 1672)=2.98, p=.02, PES=.01), and *accent* and *context* (F(4, 1672)=3.19, p=.01, PES=.01). However, the interaction for context by *listener group* (F(4, 1672)=.69, p=.60, PES=.00) was not significant. Interpretations of the significant interactions must be seen in relation to the main effects. A significant main effect on status was found for *listener group* (F(2, 1672)=6.26, p=.00, PES=.01) and *context* (F(2, 1672)=18.61, p=.00, PES=.02), but not for *accent* (F(2, 1672)=2.98, p=.051, PES=.00). Firstly, Post-hoc Tukey tests on *listener group* gave only one significant difference. The German listeners attributed significantly lower *status* to the speech samples than the Spanish listeners, and the Singaporean listeners did not differ in their *status* evaluations from the other two *listener groups*. For context, Job Pitch was evaluated significantly lower than Lecture and Audio Tour.

Subsequent univariate analyses were conducted within each *listener group* to investigate the effects of *accent* and *context* on status per *listener group*. For the German listeners, there were no significant two-way interaction effects for *accent* and *context* (F(4, 616)=2.18, p=.07, PES=.01), and no main effect for *accent* (F(2, 616)=4.16, p=.10, PES=.01). However, there was a main effect for *context* (F(2, 541)=4.16, p=.02, PES=.02). Post-hoc comparisons (HSD) showed significantly lower *status* for the speaker giving the Job Pitch, compared to Lecture and Audio Tour. Interestingly, the German listeners who had indicated when they listened to Dutch-accented English that this was a non-native accent (m=3.54; SD=.79) ascribed significantly lower status to the speaker compared to listeners who indicated that the Dutch-accented English was native (m=4.04; SD=.53) (F(1, 216)=4.34, p=.04, PES=.02).

For Singaporean listeners, there was a significant two-way interaction effect for *accent* and *context* (F(4, 541)=6.46, p=.00, PES=.05), and a main effect for *context* (F(2, 541)=4.16, p=.02, PES=.02), but there was no main effect for *accent* (F(2, 541)=1.23, p=.29, PES=.01). Post-hoc comparisons (HSD) showed significantly lower *status* for Job Pitch (m=3.68; SD=.62) compared to the Lecture (m=3.88; SD=.61) but not the Audio Tour (m=3.79; SD=.65). The two-way interaction for *accent* and *context* effect can be attributed to the different judgements on the Audio Tour for the American English accent; Audio Tour received lower scores than the other two contexts.

For the Spanish listeners, there was a significant two-way interaction effect for *accent* and *context* (F(4, 540)=3.34, p=.01, PES=.03), and a main effect for *context* (F(2, 540)=10.21, p=.00, PES=.04) and accent (F(2, 540)=5.40, p=.01, PES=.02). Post-hoc comparisons (HSD) showed significantly lower *status* for American English compared to Dutch English and British English. For *context*, post-hoc comparisons (HSD) showed significantly lower *status* for the speaker giving the Job Pitch compared to the Lecture and the Audio Tour. The two-way interaction effect for *accent* and *context* can be attributed to the absence of context differences in relation to the American English accent (Tukey's HSD).

However, in the Spanish *listener group* who listened to British English, the Job Pitch (m=3.59; SD=.72) yielded the lowest speaker status compared to the Audio Tour (m=4.05; SD=.60) and Lecture (m=4.13; SD=.72), which was similar to the results for the Spanish listener group listening to Dutch English (Job Pitch m=3.72, SD=.74; Audio Tour m=4.03, SD=.59; Lecture m=4.06, SD=.54). Finally, the Spanish listeners' evaluations of the nativeness of the standard British English *accent* and *context* showed a significant two-way interaction effect (F(2, 174)=3.53, p=.03, PES=.04). Spanish listeners who believed the standard British English accent was non-native (m=3.91; SD=.62) in the Job Pitch context ascribed significantly higher *status* to the

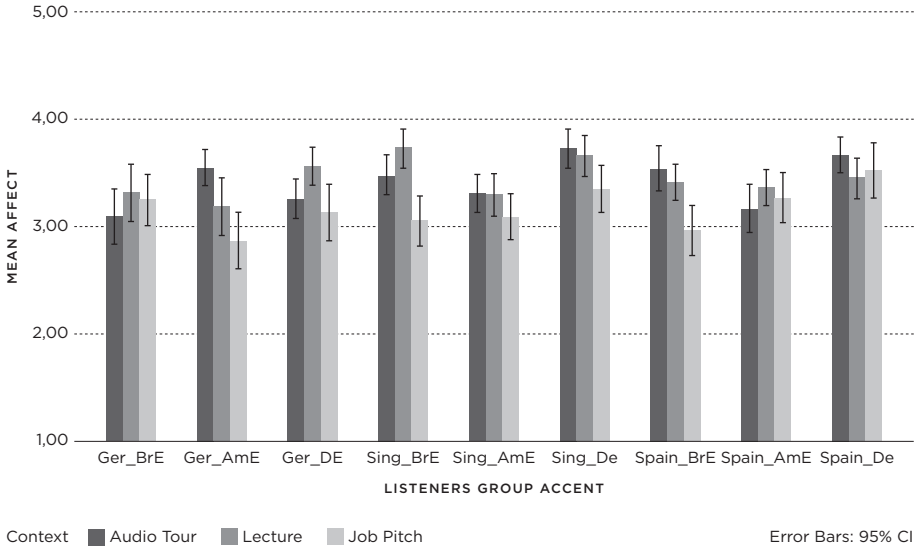
speaker compared to listeners who indicated that the British English accent was native (m=3.46; SD=.72).

5.3.8 Affect

The results for *affect* per *listener group, accent* and *context* are shown in Figure 6, with 95% confidence intervals. Univariate analyses of variance were applied to investigate the effects of *listener group*, *context* and *accent* on affect and their interaction.

The three-way interaction between *listener group, accent* and *context* (F(8, 1672)=5.11, p=.00, PES=.02) was significant. None of the two-way interactions were significant: *accent* and *context* (F(4, 1672)=.87, p=.48, PES=.00), *accent* and *listener group* (F(4, 1672)=1.30, p=.27, PES=.00), and *context* and *listener group* (F(4, 1672)=1.02, p=.40, PES=.00). The three main effects were all significant: *listener group* (F(2, 1672)=5.99, p=.00, PES=.01), *accent* (F(2, 1672)=12.83, p=.00, PES=.02), and *context* (F(2, 1672)=18.69, p=.00, PES=.02). Firstly, Post-hoc Tukey tests showed that Dutch English evoked significantly higher *affect* than both British and American

Figure 6. Affect (1=negative; 3=neutral; 5=positive) for Listener groups (Germany (Ger), Spain, Singapore (Sing)) per accent (British English (BrE), American English (AmE), Dutch English (DE)) and context (Audio Tour, Lecture, Job Pitch).



English. Secondly, German listeners attributed significantly lower *affect* to the speaker than the Singaporean and the Spanish listeners. Thirdly, the *affect* towards the speakers in the Job Pitch context was significantly lower compared with the *affect* towards the speakers in the Lecture and the Audio Tour contexts.

Subsequent univariate analyses were conducted within each *listener group* to investigate the effects of *accent* and *context* on affect per *listener group*. For the German listeners, there was a significant two-way interaction effect for *accent* and *context* ($F(4, 616)=4.99$, $p=.00$, $PES=.03$), and a significant main effect for *context* ($F(2, 541)=4.48$, $p=.01$, $PES=.02$). However, there was no significant main effect for *accent* ($F(2, 616)=.87$, $p=.42$, $PES=.00$). Subsequent post-hoc comparisons (HSD) within the listeners groups showed that the Audio Tour and Lecture aroused significantly higher *affect* than the Job Pitch (the same as the main effect in the overall analysis). The two-way interaction effect can be attributed to the different judgements in the German *listener group* of the British English accent for the three contexts: the Audio Tour achieved the lowest *affect* compared to the Job Pitch and Lecture.

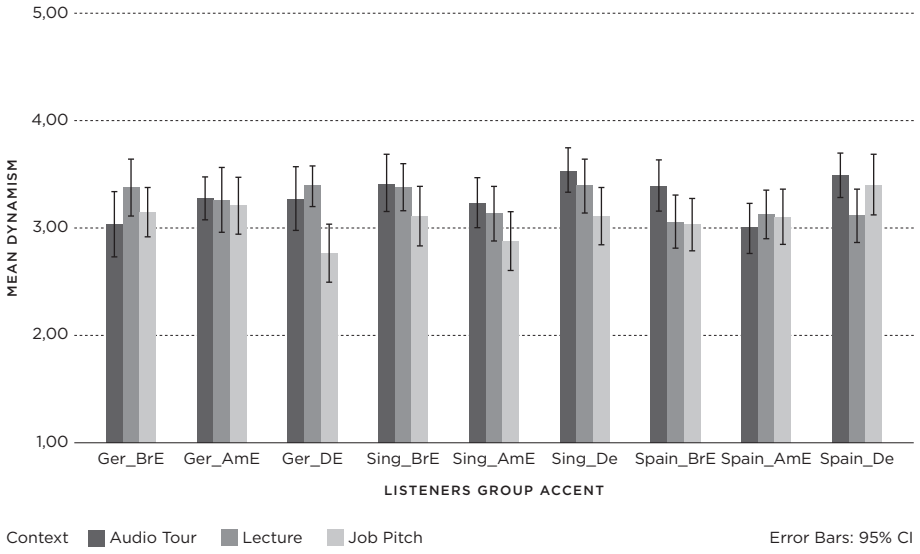
For Singaporean listeners, there was a significant two-way interaction effect for *accent* and *context* ($F(4, 541)=3.79$, $p=.01$, $PES=.03$), and main effects for *accent* ($F(2, 541)=6.67$, $p=.00$, $PES=.02$) and *context* ($F(2, 541)=3.44$, $p=.03$, $PES=.01$). Post-hoc comparisons (HSD) showed significantly higher *affect* for Dutch English ($m=3.55$; $SD=.79$) compared to American English ($m=3.27$; $SD=.78$) and British English ($m=3.30$; $SD=.83$) (the same as the main effect in the overall analysis), and significantly lower *affect* for the Job Pitch ($m=3.24$; $SD=.93$) compared to the Lecture ($m=3.41$; $SD=.67$) but not the Audio Tour ($m=3.46$; $SD=.79$). The two-way interaction effect for *accent* and *context* can be attributed to the different judgements of the three contexts within one *listener group*. There were no significant differences between the three contexts for Dutch English and American English.

For Spanish listeners, the two-way interaction effect for *accent* and *context* ($F(4, 540)=1.84$, $p=.12$, $PES=.01$) was not significant. However, there were main interaction effects for *accent* ($F(2, 540)=9.66$, $p=.00$, $PES=.04$) and *context* ($F(2, 540)=14.50$, $p=.00$, $PES=.01$). Post-hoc comparisons (HSD) showed significantly lower *affect* for American English compared to Dutch English but not British English, and significantly lower *affect* for the Job Pitch compared to the Lecture and the Audio Tour (the same as the main effect in the overall analysis).

5.3.9 Dynamism

The results for *dynamism* in relation to *listener group*, accent and context are shown in Figure 7, with 95% confidence intervals.

Figure 7. Dynamism means (1=negative; 3=neutral; 5=positive) for Listener groups (Germany (Ger), Singapore (Sing), Spain,) per accent (British English (BrE), American English (AmE), Dutch English (DE)) and context (Audio Tour, Lecture, Job Pitch).



Univariate analyses of variance showed that the three-way interaction between *listener group, accent,* and *context* (F(8, 1672)=1.76, p=.08, PES=.01) and the two-way interaction between *accent* and *context* (F(4, 1672)=.70, p=.59, PES=.00) was not significant. However, the two-way interactions between, *accent* and *listener group* (F(4, 1672)=2.59, p=.04, PES=.01), and *context* and *listener group* (F(4, 1672)=3.00, p=.02, PES=.01) were significant. The interpretations of these interactions have to be seen in relation to their main effects, of which there was one for *context* (F(2, 1672)=6.92, p=.00, PES=.01) on *dynamism*, but not for *listener group* (F(2, 1672)=.50, p=.61, PES=.00) and accent (F(2, 1672)=2.90, p=.06, PES=.00). Post-hoc Tukey tests showed that the Job Pitch evoked significantly lower *dynamism* compared with the Audio Tour and the Lecture.

Subsequent univariate analyses were conducted within each *listener group* to investigate the effects of *accent* and *context* on *dynamism* per *listener group*. For German listeners, there was no significant two-way interaction effect for *accent* and *context* ($F(4, 616)=1.99$, $p=.10$, $PES=.01$), and no main effect for *accent* ($F(2, 541)=.53$, $p=.59$, $PES=.00$). However, there was a main effect for *context* ($F(2, 616)=4.01$, $p=.02$, $PES=.01$). Subsequent post-hoc comparisons (HSD) showed that, for the German listeners, the Job Pitch evoked significantly lower *dynamism* than the Lecture, but did not differ significantly from the Audio Tour. In addition, the German listeners who had indicated that Dutch-accented English was a non-native accent ($m=2.98$; $SD=1.03$) ascribed significantly lower *dynamism* to the speaker compared to listeners who indicated that Dutch-accented English was a native accent ($m=3.71$; $SD=.84$) ($F(1, 216)=7.80$, $p=.01$, $PES=.04$). Further testing showed that when the German listeners were able to recognize Dutch-accented English as coming from a Dutchmen the assigned *dynamism* ($m=2.84$; $SD=.90$) was significantly lower compared than when they believed the accent came from a native speaker of English ($m=3.71$; $SD=.84$) ($F(1, 216)=6.86$, $p=.00$, $PES=.06$).

For Singaporean listeners, there was no significant two-way interaction effect for *accent* and *context* ($F(4, 541)=1.97$, $p=.10$, $PES=.02$), and no main effect for *context* ($F(2, 541)=2.00$, $p=.14$, $PES=.01$). However, there was a main effect for *accent* ($F(2, 541)=3.68$, $p=.03$, $PES=.01$). Post-hoc comparisons (HSD) showed that Dutch English evoked significantly higher *dynamism* than American English but not British English.

For the Spanish listeners, there was no significant two-way interaction effect for *accent* and *context* ($F(4, 540)=.08$, $p=.99$, $PES=.00$), but main effects for *accent* ($F(2, 540)=3.99$, $p=.02$, $PES=.02$), and *context* ($F(2, 540)=6.85$, $p=.00$, $PES=.03$). Post-hoc comparisons (HSD) showed significantly lower *dynamism* for American English ($m=3.08$; $SD=.97$) compared to Dutch English ($m=3.35$; $SD=.94$) but not British English ($m=3.31$; $SD=.98$). For *context*, post-hoc comparisons (HSD) showed significantly lower *dynamism* for the Job Pitch compared to the Lecture and the Audio Tour, the same effect as we found in the overall analysis.

5.3.10 Summary table speaker evaluations

Table 4 provides an overview of the results for *speaker evaluations* for the three tested accents and contexts for all three listener groups.

Table 4. Summary results Speaker Evaluations (Status, Affect, Dynamism).

N=1699	Status	Affect	Dynamism
Listener Group (L)	Spain > Germany	Spain, Singapore > Germany	NS
Accent (A)	NS	DE > BE, AE	NS
Context (C)	LE, AT > JP	LE, AT > JP	LE, AT > JP
L x A	Spain: DE, BE > AE	Spain: DE > AE Singapore: DE > BE, AE	Spain: DE > AE Singapore: DE > AE
L x C	Germany: LE, AT > JP Spain: LE, AT > JP Singapore: LE > JP	Germany: LE, AT > JP Spain: LE, AT > JP Singapore: LE > JP	Germany: LE > JP Spain: LE, AT > JP
A x C	NS	NS	NS
L x A x C	Spain BE: LE, AT > JP Spain DE: LE, AT > JP	Germany: BE: JP, LE > AT	NS

A = Accent; AE = American English; AT = Audio Tour; BE = British English; C = Context; DE = Dutch English; JP = Job Pitch; L = Listener group; LE = Lecture; NS = Not Significant.

5.4 CONCLUSION AND DISCUSSION

5.4.1 Speech understandability

Our first research question (RQ1) was aimed at investigating whether German, Spanish, and Singaporean listeners (*listener group*) displayed different *speech understandability (intelligibility, comprehensibility, interpretability)* in response to Dutch-accented English compared to standard British and American English accents (*accent*), and whether *context* (Lecture; Audio Tour; Job Pitch) affected their responses.

Our results show that *speech understandability* was not impacted overall by main effects of *accent* or *context* (see Table 2). However, the interaction effects (see again Table 2) showed that *listener groups* were affected differently by *accent* and *context* in their *speech understandability* responses. For the German listeners, standard American English was more *intelligible* compared to standard British English, and for the Spanish listeners, American English was more intelligible than Dutch English.

These results might be explained by an increased exposure to American English for these two listeners groups, however, there is no direct evidence to support this suggestion and we do not know why the Singaporean listeners did not show the same patterns. *Context* was seen to affect *comprehensibility* and *interpretability* both within and between individual listeners groups. Spanish listeners achieved higher *comprehensibility* of the audio tour than the German listeners, and higher *interpretability* of the audio tour than the German and Singaporean listeners.

For Spanish listeners, the content of the audio tour was interpreted best, followed by the lecture and job pitch. Overall, the results for the Spanish listeners illustrate that the audio tour was the context that was most understandable, in terms of *comprehensibility* and *interpretability*. The job pitch context evoked lower *comprehensibility* and *interpretability* within and across the three listener groups, and this confirms the findings from a similar study (Nejjari et al., to be published). This particular response in the job pitch context might be explained by a lack of familiarity with such a context, as our listeners were not selected specifically on the basis of experience with job pitches for retail management positions.

We had expected (expectation 1a) that Spanish listeners would display a lower level of *speech understandability* of the three accents compared to German and Singaporean listeners. This expectation was confirmed for *intelligibility*. Expectation 1a was also confirmed for Spanish listeners compared to German listeners in terms of the *comprehensibility* and *interpretability* of the job pitch. On the one hand, these results might be an illustration of previously reported weaker language skills by Spanish listeners compared to German and Singaporean listeners (EF, 2018). On the other, they might be explained by language distance , as Spanish listeners' L1, Spanish is not part of the same West Germanic language family as German, Dutch, and English. In contrast, and perhaps surprisingly, in the audio tour context, Spanish listeners achieved higher *comprehensibility* and *interpretability* compared to German listeners. This might imply that despite Spanish listeners' reported weaker fluency and their lower general *intelligibility* scores compared to German and Singaporean listeners, they were still able to correctly comprehend and interpret a particular communication context, in our case an audio tour. In turn, this might imply that an L2 listener's English fluency need not be very high in this context for them to correctly comprehend and interpret L2 speech. In addition, these findings indicate that *comprehensibility* and *interpretability* of speech can be relatively high even though *intelligibility* is relatively low, which might imply that correctly distinguishing every word (*intelligibility*) is not a prerequisite for understanding speech content (*comprehensibility*) and the communicative purpose of that content (*interpretability*). Thus, distinguishing between

these three factors of *speech understandability* in accentedness research is worthwhile, as this approach can generate more nuanced findings, and can throw light on the relationship between – and the assumed cumulative stages in – understandability of (accented) speech.

In terms of Singaporean listeners, we had expected (expectation 1b) that they would display higher *speech understandability* than German and Spanish listeners. This expectation was not confirmed for German listeners, but it was confirmed for Spanish listeners with respect to *intelligibility*. The assumed advantage for Singaporean listeners based on their higher English language skills and assumed greater exposure to various English varieties only partly emerged from the data. This might indicate that having a high level of English combined with potential exposure to varieties of English does not automatically lead to better understanding of L1 and L2 English accents.

In the context of general discussion on the existence of an LFE speech community versus traditional SLA language learning perspectives, our results indicate that effective language learning and language use may not necessarily require 'mastering the 'target' language' (L1 language) at a high level and being exposed to different varieties of that language.

Our last expectation regarding understandability (*expectation 1c*) was that the lecture context would evoke higher *speech understandability* compared to the job pitch and audio tour communication contexts. Our assumption was that our highly educated listeners would be most familiar with the lecture context, having followed higher or tertiary education. By extension, we posited that *speech understandability* would also be higher in the lecture context because of this familiarity compared to the other two contexts. *Expectation 1c* was confirmed only partly, as different patterns emerged for the different listener groups and with respect to the individual components of *speech understandability*. For instance, for intelligibility our expectation was only confirmed within the Spanish listener group. For *comprehensibility* it was confirmed for German listeners. For Singaporean and Spanish listeners the lecture context only evoked higher *comprehensibility* compared to the job pitch (but not the audio tour).

Finally, with respect to *interpretability*, both Spanish and Singaporean listeners showed higher *interpretability* of the lecture compared to the job pitch, but not the audio tour. Overall, when the lecture evoked higher *speech understandability* compared to other contexts this was only the case within the listener groups and almost only compared to the job pitch. Interestingly, our listeners did not confirm our expectation for the *speech understandability* of the lecture compared to the audio

tour, which often achieved equal or higher *speech understandability* compared to the lecture. The audio tour speech sample focused on explaining the type of art that was on display in an art gallery (e.g. Australian aborigine art; photography collections from specific time periods) and was selected from an internationally acknowledged official English certificate listening exam to reliably assess *speech understandability*. It might have been the case that our assumption with respect to the familiarity of our listeners with the three contexts was incorrect and that our listeners were familiar with an audio tour.

The results make clear that context affects *speech understandability* in various ways. It is important to understand in more detail what factors might play a role in this context effect, and therefore, further research on this topic is required. It is important to emphasize on the other hand that overall understandability scores of our respondents were very high. Lack of understanding was not a constraint for the respondents in evaluating our speakers.

5.4.2 Speaker evaluations

Our second research question (RQ2) was aimed at investigating whether German, Spanish, and Singaporean listeners (*listener group*) displayed different *speaker evaluations* in response to Dutch-accented English compared to standard British and American English accents (*accent*), and whether *context* (Lecture; Audio Tour; Job Pitch) affected their responses.

Our results show that Dutch-accented English does not have negative effects on a speaker's *status, affect,* and *dynamism (speaker evaluations)* compared with standard British and American English. In fact, it even evokes higher *affect* compared to both L1 English accents (see Table 4). However, *listener group* did impact *speaker evaluations* for *status* and *affect*, with Spanish listeners evaluating all speakers, regardless of accent, as having higher *status* compared to German listeners, and both Singaporean and Spanish listeners assigning higher *affect* to all speakers, regardless of accent, compared to German listeners. Finally, *context* affected *speaker evaluations* in that the job pitch evoked lower *speaker evaluations* compared to the lecture and the audio tour.

We had expected (expectation 2a) that German and Spanish listeners would ascribe Dutch-accented English lower *status* compared to standard British and American English accents, but this was not the case. German listeners' *speaker evaluations* were not affected by *accent*, and Spanish listeners assigned Dutch-accented English speakers significantly more positive *status, affect,* and *dynamism* than American English speakers. This suggests that in communications between L2

English speakers, having an L2 English accent can evoke equal or even higher *speaker evaluations* than a native English accent. In addition, our listeners' responses confirm Canagarajah's (2007) assertion that the LFE speech community members will likely possess flexible language attitudes, which in turn suggests that they do not hold the traditional SLA perspective on what makes for an effective L2 English speaker.

Our second expectation with regard to *speaker evaluations* (expectation 2b) was that Singaporean listeners would not respond differently to Dutch-accented English compared to standard British and American English accents. This was indeed found to be the case for *status*. Against expectations, however, for Singaporean listeners, Dutch-accented English aroused higher *affect* and *dynamism* than standard American English, and Dutch-accented English aroused higher *affect* than standard British English. This might indicate that, to Singaporeans, having an L2 accent does not negatively affect judgements of a speaker's abilities. This, in turn, might stem from living in a multicultural and multilingual society in which accent variety is the norm, and in which being an L2 English speaker can actually be an advantage if the intention is to be liked (*affect*) or perceived as dynamic (*dynamism*). In other words, this might again be an indication that the traditional SLA perspective does not automatically apply. It also illustrates the acceptance of L2 English within a specific LFE speech community, in this case Singapore.

The third expectation (expectation 2c) was that the job pitch context would evoke lower *speaker evaluations* compared to the lecture and audio tour context. This expectation was confirmed, because for all three listener groups the *status, affect,* and *dynamism* evoked by speakers in the job pitch context was lower compared to that assigned to speakers in the lecture and audio tour contexts. German listeners assigned speakers in the lecture context higher *status, affect,* and *dynamism* compared to the job pitch. They also assigned speakers in the audio tour context higher *status* and *affect* compared to speakers in the job pitch context. Spanish listeners ascribed higher *status, affect,* and *dynamism* to speakers in the lecture and the audio tour contexts compared to the speakers in the job pitch context. Finally, Singaporean listeners ascribed speakers in the lecture context higher status and *affect* compared to the speakers in the job pitch context. These findings illustrate that, except for Singaporeans and *dynamism*, the job pitch arouses lower *speaker evaluations*. Interestingly, the German listeners who evaluated the standard British English accent assigned higher *affect* to speakers in the job pitch and audio tour contexts compared to speakers in the lecture context, which is in contrast with the general finding that German listeners assigned speakers in the lecture and audio tour contexts higher *affect* compared to the speakers in the job pitch context. This

suggests that listeners may hold specific expectations of speakers in different communication contexts, and that when these expectations are violated (Burgoon et al., 1979; Burgoon & Burgoon, 2001), listeners evaluate speakers more negatively. To what extent listener expectations with regard to accentedness in different contexts actually play a role in their perceptions of speakers deserves further investigation.

Our findings do not concur with previous research that found that both L1 and L2 speakers of English tend to assign L2 English accents lower *status* compared to L1 English accents (e.g. Dalton et al., 1997; McKenzie, 2008; Nejjari et al. 2012; Nejjari et al. to be published; He & Zhang, 2010; Ryan & Sebastian, 1980; Ryan & Bulik, 1982; Cargile 1998; Lindemann, 2002; Matsuura et al., 1994). In the present study, there was a clear tolerance of L2 English speakers as listeners towards L2 English accent, as also indicated by a smaller number of studies (e.g. Bernaisch, 2012). Our results show that L2 English speakers assign L1 and L2 English accents equal *status* (e.g. Singapore, Germany), and even assign an L2 English accent higher *status* (Spain) than an L1 English accent. In addition, an L2 English accent has been shown to actually evoke higher *affect* (Spain, Singapore) and *dynamism* (Spain, Singapore) compared to L1 English accents (confirmed by Nejjari et al., to be published, Grondelaers, van Gent, et al., 2015).

During the course of our experiment we realized that the listeners who had correctly identified the accents they listened to as native or non-native might have different responses to the accents than listeners who had not. Luckily we were able to assess this on the basis of the question on the nationality of the speaker they had listened to. Our findings showed that in general there were no significant differences between listeners who had correctly identified the accents and the listeners who had incorrectly identified the accents. However, the German listeners who had identified the Dutch-accented English as non-native English, ascribed lower *status* and *dynamism* to the speaker compared to when German listeners believed the Dutch-accented English was native English, which suggests that German listeners appear to do adhere to the traditional SLA norm, even if their general evaluations show no effect of *accent*. In addition, Spanish listeners who believed the standard British English accent was non-native in the job pitch context ascribed higher *status* to the speaker compared to listeners who had indicated that the British English accent was native. This shows again that Spanish listeners preferred the L2 over an L1 English accent, and that they do not adhere to the traditional SLA norm, even in a very specific context. These effects were not part of the main questions and expectations of this study, but yielded interesting results nonetheless that call for further research.

5.4.3 Overall conclusion and implications

In general, in an LFE speech community setting, *speech understandability* and *speaker evaluations* are not negatively impacted by Dutch-accented English, but are also not positively affected by standard British or American English accents. This means that the traditional SLA perspective no longer applies to an LFE setting, where flexible language norms might hold instead. Most 'foreign' or L2 language education is aimed at training language learners to become fluent on a 'native level', in various degrees, in a 'target' language, which typically is a standard, L1 variety of the 'target' language. In the case of English, L2 English speakers and English language learners now outnumber L1 English speakers (Crystal, 2009). Our results suggest that educators and learners of English might need to review their perceptions of what it means to be a successful learner and speaker of English, and potentially rethink educational practices. Our results show, contrary to many accentedness studies, that an L2 English accent within the LFE speech community does not have to hinder effective communication. First, an L2 English accent can be as understandable as standard L1 English accents, even if listeners are not familiar with that particular L2 English accent, which is a realistic scenario in international interactions. Second, if the intent is to evoke perceptions of high status, affect, and dynamism in listeners in the LFE speech community, being an L2 English speaker with an L2 English accent can actually be more beneficial compared to having a standard, L1 English accent. This means that for L2 English speakers accent training may not be necessary if achieving understandability and positive speaker evaluations are the main learning objectives.

Context, however, did impact *speech understandability* and *speaker evaluations* with the job pitch context almost universally achieving lower understanding and lower evaluations compared to the lecture and audio tour contexts. These results suggests that listeners might not necessarily respond to a speaker's accent, instead, they react to the communication context in which a speaker communicates, which might me connected to expectations concerning these contexts. As a result, breakdowns in communication between speakers of English with various L1 backgrounds should not solely be evaluated in terms of speakers' language skills and accentedness, but should also be analyzed in relation to the familiarity and requirements of the communication context.

5.4.4 Limitations and future research

This study has a number of limitations. First, we used only one male matched-guise speaker in our experiment. This could have led to reactions that represent a response to this particular male speaker, and to male speakers only. At the same time, the use

of the matched-guise technique contributed to the validity of our results, since they cannot be attributed to the voice characteristics of individual speakers (as might have been the case had verbal guises been used).

A second limitation was that the listeners were selected to represent people who in general would be most likely to be familiar with the three selected contexts. With hindsight, however, professionals who regularly interact in these contexts, such as HR managers in the job pitch context, would have been the optimal choice, but this was not possible for practical reasons.

Third, the accents studied are from one language family, and future studies should on the basis of linguistic relatedness select accents to see whether the accent itself or other linguistic features, such as prosody affect listeners' responses. This can help language learners as well as L1 speakers of languages understand the effects their accents may have on listeners and perhaps avoid prejudices caused by accentedness.

In terms of the assessment of *speech understandability* based in Kachru and Smith's (2008) three speech understanding components, *intelligibility, comprehensibility,* and *interpretability* yielded results that provide useful insights into the levels at which a varied group of listeners are able to understand speech. However, in order to assess *comprehensibility,* and *interpretability,* two statements on each speech sample were presented to listeners on its content (*comprehensibility*) and its communicative purpose (*interpretability*), which had to be judged as correct or incorrect. This is a very general manner to assess these components of *speech understandability*. Future research should seek more detailed measurements to further understand how these components can be studied best, and also assess how these three components are connected, with the aim to understand how *speech understandability* works and how better understanding of speech can be achieved. Additionally in terms of *speaker evaluations,* the *dynamism* questionnaire items loaded as part of a third separate factor for the German and Spanish listeners, but did not for the Singaporean listeners. This might indicate that German and Spanish listeners define a speaker's dynamism differently from Singaporean listeners. Further research on how different listener groups define a speaker's dynamism could help researchers select items that optimally study the perceptions listeners have of different speaker groups.

Finally, the data was collected via Qualtrics, an online data collection service, which is increasingly used by scientists across the globe. Qualtrics guarantees high quality data collection and it was the only way that allowed swift international data collection of a significant number of responses. We do not see this as a limitation, but as an incentive to supplement our data with accent evaluation research on a more restricted local level, with alternative ways of sampling of respondents.

5.5 SUPPORTING INFORMATION

5.5.1 Speech sample links study 4 (and 3): matched guises, filler, controls

The speech sample links below redirect to one speech sample per accent (standard British, standard American English, Dutch English) and communication context (lecture, audio tour, job pitch), per speaker (matched-guise speaker; filler speaker; control speakers). For the controls, three speech sample links are available as examples (for each accent one).

Matched-guise speaker Lecture
- standard British English https://cls.ru.nl/webexp-media/HV_BE_L.html
- standard American English https://cls.ru.nl/webexp-media/HV_AE_L.html
- Dutch English https://cls.ru.nl/webexp-media/HV_DE_L.html

Matched-guise speaker Audio tour
- standard British English https://cls.ru.nl/webexp-media/HV_BE_A.html
- standard American English https://cls.ru.nl/webexp-media/HV_AE_A.html
- Dutch English https://cls.ru.nl/webexp-media/HV_DE_A.html

Matched-guise speaker Job pitch
- standard British English https://cls.ru.nl/webexp-media/HV_BE_J.html
- standard American English https://cls.ru.nl/webexp-media/HV_AE_J.html
- Dutch English https://cls.ru.nl/webexp-media/HV_DE_J.html

Filler speaker
- standard British English https://cls.ru.nl/webexp-media/CG_BE_F.html

Examples of controls or native speakers for each accent (in the lecture context only)
- standard British English https://cls.ru.nl/webexp-media/JC_BE_L.html
- standard American English https://cls.ru.nl/webexp-media/PG_AE_L.html
- Dutch English https://cls.ru.nl/webexp-media/CJ_DE_L.html

5.5.2 Country of origin of speakers estimates per listener group and per accent in rounded %

	Standard British English	Standard American English	Dutch-accented English
Germany (n = 617)	47.4% correct Incorrect: 31.4% U.S.; 13.7% Australia/ New Zealand; 4.0% (Western) Europe/L2; 2.3% German; 1.1% Africa	47.5% correct Incorrect: 31.9% U.K.; 12.7% Australia/ New Zealand 4.9% (Western) Europe/L2; 2.5% German; 0.5% Africa	40.1% correct Incorrect: 19.8% U.K.; 24.0% Europe, France, Belgium/L2/the West; 6.0% U.S.; 6.0% Germany; 4.1% Asia; 0.5% Australia/ New Zealand
Spain (n = 540)	57.4% correct Incorrect: 22.8% U.S.; 14.2% Australia / New Zealand; 3.6% Europe/ German/L2; 0.6%Netherlands; 0.6% South America; 0.6%Spain	42.3% correct Incorrect: 36.2% U.K.; 14.1% Australia / New Zealand; 5.0% Europe/ German/L2; 0.6% Netherlands; 0.6% Asia; 0.6% Spain; 0.6% South America	1.7% correct; Incorrect: 35.6% Europe, France, Germany, Russia, Belgium/L2; 28.2% U.K.; 16.7 % U.S.; 11.5% Australia/ New Zealand; 4.0% Asia; 1.7% Spain; 0.6% Africa.
Singapore (n = 542)	41.0% correct Incorrect: 28.1% U.S.; 17.5% Australia/ New Zealand 7.9% Europe 2.8% Asia 2.2% Netherlands; 0.5% L2	48.5% correct Incorrect: 23.4% U.K.; 15.0% Australia / New Zealand; 8.4% Europe; 2.4% Asia; 0.6% Netherlands; 1.2% South America; 0.5% Africa	1.7% correct Incorrect: 54.6% Europe, France, Germany, Russia, Italy; 13.2% U.K.; 11.5 % U.S.; 9.2% Singapore/Asia; 6.9% Australia/ New Zealand; 2.2% the West/western countries/L2; 0.6% Africa

5.5.3 Germany speech understandability and speaker evaluations per accent (British English, American English, Dutch English; 1=negative; 3=neutral; 5=positive) and context (Lecture, Audio Tour, Job Pitch)

	Speech understandability			Speaker evaluations		
Accent, Context	Intelligibility Mean (SD)	Comprehensibility % correct	Interpretability % correct	Status Mean (SD)	Affect Mean (SD)	Dynamism Mean (SD)
Dutch English Lecture[a]	7.22(3.61) n=114	91.2% n=104	76.3% n=87	3.83(.70) n=114	3.55(94) n=114	3.39(1.12) n=114
Dutch English Audio Tour[a]	7.68(3.64) n=59	72.9% n=43	78% n=46	3.75(.60) n=59	3.27(.69) n=59	3.29(.99) n=59
Dutch English Job Pitch[b]	7.37(3.80) n=52	80.8% n=42	75% n=39	3.34(.94) n=52	3.13(.93) n=52	2.77(.98) n=52
British English Lecture	7.67(3.93) n=49	93.9% n=46	89.8% n=44	3.91(.78) n=49	3.35(.93) n=49	3.37(.92) n=49
British English Audio Tour	7.38(3.91) n=48	77.1% n=37	79.2% n=38	3.71(.72) n=48	3.09(.89) n=48	3.04(1.04) n=48
British English Job Pitch	6.31(4.11) n=87	78.2% n=68	70.1% n=61	3.72(.73) n=87	3.22(1.07) n=87	3.16(1.06) n=87
American English Lecture	8.34(.25) n=53	84.9% n=45	86.8% n=46	3.81(.67) n=53	3.18(.94) n=53	3.26(1.04) n=53
American English Audio Tour	8.01(4.01) n=102	81.4% n=83	76.5% n=78	3.84(.71) n=102	3.56(.87) n=102	3.28(1.00) n=102
American English Job Pitch	7.83(3.67) n=53	84.9% n=45	83% n=44	3.66(.56) n=53	2.89(.93) n=53	3.23(.94) n=53

[a]max. 11 words; [b]max. 12 words intelligible. N=617; n=number of listeners per accent and context.

5.5.4 Singapore speech understandability and speaker evaluations per accent (British English, American English, Dutch English; 1=negative; 3=neutral; 5=positive) and context (Lecture, Audio Tour, Job Pitch)

	Speech understandability			Speaker evaluations		
Accent, Context	Intelligibility Mean (SD)	Comprehensibility % correct	Interpretability % correct	Status Mean (SD)	Affect Mean (SD)	Dynamism Mean (SD)
Dutch English Lecture[a]	6.12(3.84) n=59	83.1% n=49	86.4% n=51	3.71(.60) n=59	3.45(.72) n=59	3.12(.82) n=59
Dutch English Audio Tour[a]	6.29(3.55) n=63	85.7% n=54	69.8% n=44	3.87(.50) n=59	3.67(.65) n=59	3.49(.95) n=59
Dutch English Job Pitch[b]	6.66(4.12) n=58	65.5% n=38	69% n=40	3.78(.64) n=58	3.52(.96) n=58	3.41(1.06) n=58
British English Lecture	7.20(4.10) n=64	89.1% n=57	82.3% n=53	3.98(.63) n=64	3.41(.66) n=64	3.06(.97) n=64
British English Audio Tour	7.07(3.77) n=60	78.3% n=47	83.3% n=50	3.96(.66) n=60	3.54(.81) n=60	3.40(.92) n=60
British English Job Pitch	7.68(3.97) n=63	68.3% n=43	71.4% n=45	3.55(.68) n=63	2.96(.91) n=63	3.03(.95) n=63
American English Lecture	7.19(4.31) n=62	85.5% n=53	83.9% n=52	3.94(.57) n=62	3.36(.64) n=62	3.13(.88) n=62
American English Audio Tour	6.60(4.40) n=58	81% n=47	74.1% n=43	3.52(.69) n=58	3.16(.84) n=58	3.00(.88) n=58
American English Job Pitch	6.69(4.66) n=55	74.5% N=41	72.7% n=40	3.73(.51) n=55	3.26(.85) n=55	3.11(.92) n=55

[a]max. 11 words; [b]max. 12 words intelligible. N=542; n=number of listeners per accent and context.

5.5.5 Speech understandability and speaker evaluations per accent (British English, American English, Dutch English; 1 = negative; 3 = neutral; 5 = positive) and context (Lecture, Audio Tour, Job Pitch)

	Speech understandability			Speaker evaluations		
Accent, Context	Intelligibility Mean (SD)	Comprehensibility % correct	Interpretability % correct	Status Mean (SD)	Affect Mean (SD)	Dynamism Mean (SD)
Dutch English Lecture[a]	5.23(3.10) n=61	98.4% n=60	90.2% n=55	4.06(.54) n=61	3.65(.74) n=61	3.39(.95) n=61
Dutch English Audio Tour[a]	7.26(3.16) n=61	83.6% n=51	83.6% n=51	4.03(.59) n=61	3.73(.69) n=61	3.54(.79) n=61
Dutch English Job Pitch[b]	5.33(2.91) n=60	76.7% n=46	65% n=39	3.72(.74) n=60	3.35(.84) n=60	3.12(1.03) n=60
British English Lecture	6.12(2.78) n=60	88.3% n=53	90% n=54	4.13(.759) n=60	3.73(.67) n=60	3.38(.82) n=60
British English Audio Tour	7.07(2.73) n=60	76.7% n=46	78.3% n=47	4.05(.60) n=60	3.48(.70) n=60	3.42(1.01) n=60
British English Job Pitch	5.37(3.32) n=60	70% n=42	61.7% n=37	3.59(.58) n=60	3.05(.90) n=60	3.12(1.06) n=60
American English Lecture	7.09(3.47) n=59	91.5% n=54	93.2% n=55	3.69(.65) n=60	3.29(.75) n=60	3.13(.96) n=60
American English Audio Tour	6.98(3.10) n=60	93.3% n=56	83.3% n=50	3.77(.72) n=59	3.31(.67) n=59	3.24(.90) n=59
American English Job Pitch	6.37(3.16) n=60	78% n=46	59.3% n=35	3.74(.68) n=59	3.08(.83) n=59	2.88(1.04) n=59

[a]max. 11 words; [b]max. 12 words intelligible. N = 540; n = number of listeners per accent and context.

6
CONCLUSION

The main objectives of this thesis were (1) to investigate L2 and L1 English listeners' responses to Dutch-accented English compared to L1 standard British and American English accents, in terms of *speech understandability* and *speaker evaluations*, and (2) to understand whether responses to accents are relative to the communication context in which they are heard. This final chapter discusses the results of the four studies we carried out and considers implications for L2 English learning. Section 6.1 discusses the studies' main results in terms of *speech understandability, speaker evaluations,* and *communication context*. Section 6.2 focuses on what the studies' results mean for the traditional second language acquisition (SLA) perspective and the concept of a lingua franca English (LFE) speech community. Section 6.3 outlines the implications of our research for English language learning and teaching, and section 6.4 offers recommendations for future research and methodological considerations.

6.1 MAIN RESULTS: SPEECH UNDERSTANDABILITY, SPEAKER EVALUATIONS, AND COMMUNICATION CONTEXT

6.1.1 Understanding Dutch-accented English: speech understandability

The findings from the three studies reported in Chapters 2, 4, and 5 show that for the Dutch, German, Singaporean, and Spanish listener groups (L2 English listener groups), Dutch-accented L2 English, compared to standard British and American L1 English accents, does not hinder *speech understandability* in terms of intelligibility (literally distinguishing words and phrases), comprehensibility (understanding the meaning and content of the speech samples), and interpretability (understanding the communicative intentions of the speaker). However, for the British listener group, Dutch-accented English does hinder intelligibility and comprehensibility, at least compared to a standard British English accent.

It is important to note that, as discussed in Chapter 1, study 1 differed from studies 3 and 4, in that study 1 focused on two degrees of Dutch-accented English (slight, moderate) compared to a standard British English accent. In addition, it employed a slightly different operationalization of the *speech understandability* dimensions comprehensibility and interpretability. The *speech understandability* for all three accents in study 1 was never low for the British listeners (see sections 2.3.2 – 2.3.4). However, in contrast to the four L2 English listener groups, the British listeners found standard British English more intelligible and comprehensible than a Dutch English accent (both slight and moderate accents). Interpretability was not affected by either of the two degrees of Dutch-accented English, compared to a standard British English. Familiarity with a Dutch-English accent was an independent variable in study 1, and was found to play a role, in that the British listeners familiar with Dutch-accented English found it more intelligible than the British listeners who were not familiar with Dutch-accented English.

Similar to study 1, the *speech understandability* of the accents tested in studies 3 and 4 was never poor, for any of the four L2 English listener groups (Chapters 4 and 5), regardless of whether they were familiar with the tested accents or not (sections: 4.3.1, 4.5.2, 5.3.1-5.3.5, 5.5.3-5.5.5). Compared to the other L2 English listeners, the Dutch listeners in studies 3 and 4 achieved the highest correct accent identification scores for all three English accents, suggesting that they had the highest familiarity with them, followed by the German, Singaporean and Spanish listeners (see sections 4.2.3, 5.5.2). It was shown that for the Dutch listeners, all of whom were familiar with Dutch-accented English, *speech understandability* for this accent was the highest for all three dimensions. The findings with regard to intelligibility for the Dutch

listeners correspond with Wang and van Heuven (2007) and Bent and Bradlow (2003) who found that vowel, consonant, and sentence recognition (or intelligibility) is higher when listeners share a native language with the speaker, a mechanism they termed interlanguage speech intelligibility benefit (Bent & Bradlow, 2003). For the German, Singaporean, and Spanish listener groups Dutch-accented English, compared to standard British and American English accents, also did not hinder *speech understandability* with regard to any of the three dimensions, but there were differences between the listener groups that are in line with the extent to which they recognized it as a non-native English accent. For example, the three tested accents were less intelligible for the Spanish listeners than for the Singaporean and German listeners, while for the Singaporean listeners, they were less intelligible than for the German listeners.

The findings of the three studies reported in chapters 2, 4, and 5 suggest that compared to L1 English listeners, L2 English listeners are generally not at a disadvantage in understanding L1 and L2 accent varieties of English, and may even have an advantage over L1 English listeners at the level of intelligibility of L2 English accents. Perhaps this is because learners of English, who are most likely to use English with other L2 English speakers (e.g. in an LFE context), are exposed to multiple L2 English accents. This exposure may promote a language processing flexibility that allows L2 English listeners to more easily recognize sound and word patterns that are non-native.

6.1.2 Speaker evaluations: status, affect, and dynamism

The findings from the three studies reported in Chapters 2, 4, and 5 show that for L2 English listeners from Singapore and Spain, a Dutch-accented L2 English, compared to standard British and American L1 English accents, does not negatively impact *speaker evaluations* for status, affect, and dynamism (sections 5.3.6–5.3.10). These results disconfirm the findings from previous research that suggest that L2 English accents can negatively impact perceptions of speakers (e.g. Dalton-Puffer, Kaltenbroek & Smit, 1997; Zhang, 2010; Ryan & Sebastian, 1980; Ryan & Bulik, 1982; Cargile, 1997; Cargile & Giles, 1998; Lindemann, 2002; McKenzie, 2008; Matsuura, Chiba & Yamamoto, 1994).

German listeners were generally very similar to the Singaporean and Spanish listeners in their responses to the three accents with regard to the affect of speakers, but differed in their status and dynamism evaluations. German listeners who had indicated that the Dutch-accented English was a non-native accent, ascribed significantly lower status and dynamism to the Dutch English speaker compared to German listeners who indicated that the Dutch-accented English was native (sections

5.3.7 , 5.3.9, 5.3.10). For Dutch listeners (section 4.3.2), Dutch-accented English, compared to standard British and American English accents, negatively affected evaluations of status, but not affect and dynamism. The study with British listeners (Chapter 2) did not include dynamism evaluations and compared two degrees of Dutch-accented English (slight and moderate) to standard British English. It found that the British listeners generally evaluated the status of Dutch-accented English speakers (slight and moderate accents) and the affect of moderate Dutch-accented English speakers significantly more negatively compared to a standard British English speakers (section 2.3.1).

In studies 2, 4 and 5, both the Dutch and British listeners, and German listeners who had indicated that the Dutch-accented English was a non-native accent,were found to assign lower status to Dutch-accented English speakers, compared with This shows that British, Dutch, and some German listeners in the present study associate a non-native accent with lower social status than a native accent. Degree of accentedness (of Dutch English) was only included in study 1 with British listeners, where it did not impact speakers' status. It might be the case that, to the British listeners, the fact that the accent is non-native matters more than the degree of accentedness. British and Dutch listeners associate non-nativeness with lower status. This association might indicate that British (study 1), Dutch (study 3), and some German listeners (study 4), unlike Singaporean and Spanish listeners (study 4), share a traditional perspective on successful SLA: the best language speaker is an L1 speaker, and therefore, second language learning is most effective when an L2 speaker can speak a 'target' language like an L1 speaker (see also section 6.2).

With regard to the affect for a speaker, the Dutch, German, Singaporean, and Spanish listeners generally evaluated Dutch-accented English favourably. They did not evaluate speakers with a Dutch English accent as evoking lower affect than speakers with standard L1 English accents (sections 4.3.2, 5.3.8 and 5.3.10). In fact, Dutch-accented English aroused higher affect than standard American English in both Singaporean and Spanish listeners, and compared to standard British English in Singaporean listeners (sections 5.3.8 and 5.3.10). This seems to imply that, if L2 English speakers want to be liked within a speech community of L2 English speakers, it pays to sound non-native. The assumption is that shared non-nativeness promotes solidarity, and by extension, the affect associated with speakers that form part of the community.

For British listeners, the affect for slightly Dutch-accented English speakers and standard British English speakers was comparable (section 2.3.1). A moderate Dutch English accent, however, aroused significantly lower affect for the speaker in British

listeners, suggesting that non-nativeness does not necessarily impact affect negatively for British listeners, provided it is not too non-native. Research by van den Doel (2006) might offer an explanation for the British listeners' different affect evaluations of slight and moderate Dutch English accents. Van den Doel (2006) developed a hierarchy of pronunciation errors in the English of Dutch speakers of English, based on the judgements of a variety of L1 English speakers. He shows that some pronunciation errors, such as /æ ~ e/ confusion, were ranked very highly as significant errors, but that the selected L1 English listener groups also differed from one another in their ranking of certain errors. The different degree of affect assigned to a speaker with a slight Dutch accent and a speaker with a moderate Dutch accent by the British listeners in study 1 could be explained by such a pronunciation error hierarchy. When certain pronunciation errors are judged as too deviant from a standard British English accent, affect decreases for these British listeners.

Similar to the results for affect, the Dutch, German, Singaporean, and Spanish listeners did not assign lower dynamism to speakers of Dutch-acccented English than to speakers of standard British and American English (sections 4.3.2, 5.3.9 and 5.3.10). In fact, both the Singaporean and Spanish listeners assigned Dutch-accented English speakers higher dynamism than standard American English speakers. Again, this seems to suggest that, to L2 English listeners, a speaker's dynamism can be influenced by accent, and that non-nativeness can even impact a speaker's dynamism positively.

In studies 3 and 4, the results for affect and dynamism appear to indicate that L2 English listeners have sympathy with L2 English speakers, at times more so than with L1 English speakers, and that some L2 English listeners view L2 English speakers as more dynamic than L1 English speakers. This might be because L2 English speakers are likely to have experienced for themselves that language learning can be very challenging and requires effort and perseverance, which makes L2 English listeners believe L2 English speakers are dynamic. This means that a non-native accent is not necessarily a disadvantage when it comes to how a speaker is evaluated, especially not in communications with fellow L2 English listeners who have a different L1 background.

6.1.3 Correlations

Correlations were calculated to explore the potential relationship between *speech understandability ratings* and *speaker evaluations* (sections 2.3.5 and 4.3.3). In study 1, involving British listeners, there was a positive correlation between intelligibility and comprehensibility. When a speaker is more intelligible, that person is considered

more comprehensible as well. Positive correlations were also found for status with affect, intelligibility, and comprehensibility: when a speaker is assigned higher status they are also regarded as more intelligible and comprehensible, and evoke higher affect. Finally, affect was positively correlated with comprehensibility, meaning that when a speaker evokes more affect, that speaker will also be more comprehensible.

In study 3, involving Dutch listeners, we observed correlations in the opposite direction, with affect showing a weak, but significantly negative correlation with comprehensibility. This means that, as speakers arouse higher affect, the comprehensibility of their speech decreases. Overall, these results might indicate that the relationship between the tested variables and their individual dimensions can differ per listener group, but future research should be conducted to further explore this assumption (see also section 6.4). A problem we encountered as we progressed through the studies is the ceiling effects found for *speech understandability* in study 3, which may have had a suppressive effect on correlation sizes. We therefore decided not to include correlation calculations in study 4, which was already fairly lengthy, and again, acknowledge that we should explore correlations further in future studies.

6.1.4 Communication context in accentedness studies

Like study 1, which investigated accentedness in a single international business communication context, accentedness studies have tended to focus on one particular communication context in which accentedness may be relevant, such as sales or education (e.g. Hendriks et al., 2016, 2018; Bolton & Kuteeva, 2012; Hellekjaer, 2010; Tsalikis, DeShields & LaTour, 1991). After study 1 we began to wonder whether responses to accents might differ relative to the communication context in which they are heard, and investigating this question became the second overall aim of this thesis.

Studies that have tested accents in more than one communication context are rare. One such study (Cargile, 1997) showed that responses to accents in different communication contexts can indeed vary. As discussed in Chapters 4 and 5, when people expect certain non-verbal and verbal behaviors from others in a specific communication context, violations of expectations may impact people's reactions to, and perceptions of, others. Such outcomes match Expectancy Violation Theory, which is focused on non-verbal behavior (see Burgoon, Stacks & Woodall, 1979; Burgoon & Le Poire, 1993, Burgoon & Burgoon, 2001).

Within an L2 English speakers' speech community in particular, where different English speaker groups might have different expectations in terms of someone's English accent, responses to English accentedness can be expected to be context-

dependent. For example, within the present study, the Singaporean listeners are from a society in which English is widely used as a communication tool to communicate across ethnic boundaries in educational, governmental, and professional contexts (section 5.1.2). As a result, it could be the case that Singaporeans have different expectations about English and what is considered acceptable English language use than people from societies in which English is also an important L2, but in which English does not function as an important tool to communicate across ethnic boundaries. In order to investigate the potential effects of *communication context*, in studies 3 and 4, responses to accentedness were studied in three contexts that were thought to be relevant to the highly educated L2 English listeners: (1) higher education (a higher education marketing lecture), (2) tourism (an art gallery audio tour), and (3) business (a sales pitch, and a job pitch for a management position).

The conclusion that can be drawn overall with regard to responses to accentedness in different contexts is that the listener groups in studies 3 and 4 responded more strongly to context than accent, and that apart from a few exceptions, there was no interaction between the two. The main pattern was that communication context impacts responses to speech in general. For instance, overall, in studies 3 and 4, a speaker in the lecture context was better understood and evoked more positive speaker evaluations than a speaker in the job pitch context, and the speaker in the audio tour context was equally well understood and evaluated as the speaker in the lecture, regardless of accent, and for all L2 English listener groups. This might be explained by the fact that the listener groups in the present study were all highly educated and had reported at least average English fluency. They are therefore likely to have attended lectures in a higher education setting and to have been exposed to 'high culture' activities, such as a supervised (audio) tour of a museum or art gallery.

The addition of communication context as a variable in our later studies has allowed us to view responses to accents across various relevant domains of language use. The results indicate that communication context matters and impacts responses to speakers and understandability.

6.2 THE SLA PERSPECTIVE AND THE LFE SPEECH COMMUNITY

Due to the status of English as a global language, used mostly by L2 English speakers as a lingua franca, linguists have questioned the traditional second language acquisition (SLA) perspective, as explained in Chapters 1 and 5. The criticism is aimed at the traditional assumption that SLA involves the learning of an L1 language variety by

non-native speakers whose own native language often 'interferes' in the successful acquisition of the L1 variety in question, or the 'target language'. According to this traditional perspective, successful SLA would entail mastering an L1 language variety with limited interference by the learner's L1. This is said to facilitate better mutual understanding between English speakers with various language backgrounds (see Gerritsen et al., 2016 for a similar discussion in the context of Dutch English). The reason why this perspective is criticized is that it assumes that L2 speakers of languages should try to be as native as possible in order to become successful communicators (Firth & Wagner, 1998; Canagarajah, 2007), which is hardly possible (e.g. Lenneberg, 1967; Bongaerts, van Summeren, Planken & Schils, 1997; Friedman & Rusou, 2015). In addition, the traditional SLA viewpoint might not necessarily be universally relevant and is said to ignore the social and interactive nature of languages. Canagarajah (2007) claims that the status of English as a global lingua franca and the resulting Lingua Franca English (LFE) speech community show that there is no need to be as native as an L1 English speaker. According to him, this is because in the LFE speech community, which is diverse in nature, English use activates a set of language norms (e.g. which accent varieties are considered prestigious) and communication strategies (e.g. adaptation of grammar, lexicon, pragmatic conventions to suit interactions) that facilitate successful communications between speakers who are sufficiently proficient in English, but who are not L1 English speakers. Canagarajah stresses that LFE should not be viewed as a separate language system. Instead, he claims that LFE comes into being when LFE speech community members communicate with each other, and changes according to the communication requirements of each individual interaction.

The results of our four studies show that British and Dutch listeners, and German listeners who are aware they are evaluating a non-native English accent, have attitudes towards Dutch-accented English that fit the traditional SLA perspective, since they assigned Dutch-accented English lower status, compared to the L1 English accents, even when they had no difficulties in understanding the Dutch-accented English. This was particularly interesting in the case of the Dutch listener group (section 4.3.2). They correctly identified all three accents and thus knew who they were evaluating, and they experienced no difficulties whatsoever in understanding these accents. Yet, they still 'preferred' standard British and American English to their 'own' accent in English at least where the status a speaker evokes was concerned. This negative judgement of Dutch-accented English compared to L1 English accents seems to reflect a norm that defines a competent, cultured, and intelligent L2 English speaker as someone who sounds as native as possible.

Interestingly, the Singaporean and Spanish listener groups from study 4 displayed no negative evaluations of Dutch-accented English at all. In terms of evaluations, the German, Singaporean, and Spanish listeners were not negatively impacted by a speaker's accent. In some cases they even 'preferred' the Dutch-accented English to the L1 English accents (sections 5.3.6-5.3.10). Contrary to the findings for the Dutch listeners then, the findings for the Singaporean, Spanish, and to some extent the German listeners, appear to support Canagarajah's claim that L2 English speakers – amongst themselves – have developed their own speech community, which has its own perspective on what a high status, likeable and dynamic English speaker sounds like. Therefore, it would seem L2 English speakers are not necessarily at a disadvantage when they communicate with other L2 English speakers, in terms of how they are understood and evaluated.

Overall, our results imply that the traditional views on successful L2 or second language acquisition (SLA) should be reconsidered. Many accentedness studies focus on L1 speakers' evaluations, and study 1 with the British listeners confirmed the findings of earlier studies. However, the results from study 4 in particular showed that L2 English listeners can vary in their speaker evaluations of an L2 English accent, but that an L2 English accent can also be evaluated more positively than L1 English accents. In addition, L2 English listeners understand L2 and L1 English speakers equally well. These results suggest that successful L2 English acquisition does not need to focus on obtaining an L1 English accent.

6.3 IMPLICATIONS FOR ENGLISH LANGUAGE TEACHING AND LEARNING

The four studies have shown that, apart from British listeners, all other listener groups had no difficulties with the L2 and L1 English accents under study in terms of understanding, even when they were not familiar with these accents. An L2 English accent positively impacted perceptions of affect and dynamism and only impacted evaluations of speakers in terms of status negatively.

In contrast to accent however, communication context proved to significantly impact how speakers were perceived. This means that teachers of English should take into consideration in their pronunciation training in which contexts and situations their students have to communicate and focus their pronunciation trainings on these contexts. L2 English speakers have to decide what they believe is the most helpful strategy in their individual situation, and in which communication contexts they will use English. If it is certain that most English use will be mainly with other L2 English

speakers, obtaining a native-like accent will not be necessary. However, when this is not the case, it might be helpful to sound as native as possible, in order to best guarantee mutual understanding and positive perceptions, as has been recommended by others (e.g. van Oostendorp, 2002). It is true that these results apply to Dutch-accented English only, but the results can also be interpreted as a persuasive example of how non-native accentedness does not automatically lead to negative perceptions or decreased understanding. Therefore, L2 English speakers, but also L1 English speakers and English teachers, can on the basis of these results become more aware of their potential prejudice towards L2 English accents and language varieties in general and attempt to refrain from negative evaluations of L2 English accents in general.

6.4 FUTURE RESEARCH AND METHODOLOGICAL CONSIDERATIONS

The four studies reported on reflect a research development process that started with the first study a few years before the final three studies were conducted. The experience gained in study 1 inspired the design and methodology of the following three studies. For example, in study 1 *speech understandability* was measured by employing Kachru and Smith's (2008) three dimensions of understanding speech: (1) intelligibility, which was defined as a listener's ability to distinguish words and phrases measured by orthographic transcription, (2) comprehensibility, which was defined as a listener's ability to understand the meaning of words and phrases and intentions in the proper context and measured by asking them to paraphrase speech and select key words to ascertain whether they had understood an utterance's meaning, and (3) interpretability, which was defined as a listener's ability to understand the purpose of a communicative act and measured by asking them what they believed the purpose was of a speech sample excerpt.

As indicated in Chapters 4 and 5, this definition of *speech understandability* was maintained for intelligibility in in studies 3 and 4. Comprehensibility and interpretability were defined similarly, but measured differently in studies 3 and 4. In studies 3 and 4, statements on the content (comprehensibility) and the communicative purpose of the speech content (interpretability) were posed which the listeners could indicate were correct or incorrect. This might have increased the listeners' ability to correctly guess the correct answer to the question. To our knowledge, this has been the first attempt of measuring speech understandability in this manner, and future research could further investigate how to best define *speech understandability* and further optimize its definition and operationalization.

Studies 1, 3, and 4 measured *speaker evaluations*. In studies 3 and 4 dynamism was added on the basis of research that was published after study 1 was conducted (Grondelaers and van Hout, 2015). Dynamism proved to offer a relevant contrast with status in Grondelaers and van Hout (2015). They showed that non-standard language use can impact a person's perceived status, but not dynamism. In study 3, the assessment of the three accents for status, affect, and dynamism was successful; however, in study 4 the factor loadings for status, affect, and dynamism for listener group differed from the Dutch listeners in study 3. For instance, 'confident' was an item that was selected to measure dynamism, which it did in the Dutch listener group, but this was for German, Spanish, and Singaporean listeners considered part of status. In addition, Singaporean listeners made no distinction between dynamism and both status and affect. It was decided to analyze the responses to the dynamism item 'energetic' as representing dynamism, since the German and Spanish listeners defined it as separate from status and affect. The questionnaires were translated by translators and language experts into the listeners' native languages to make sure that items had the same meaning in the translated questionnaires as they did in the original questionnaire. After the translations were finished, the translations were informally discussed with the translators by the researcher to check whether the correct translation had been achieved. Even though most factor loadings were the same for all listener groups, the different responses to the individual items might have been a result of cultural differences in terms of the elements that constitute status, affect, and dynamism. Future research should use back translations and pretests to determine whether evaluative items suit the defined listener groups, and design research tools to optimally assess speaker evaluations.

In order to investigate representative accents, study 1 employed a verbal-guise technique, which entailed selecting native speakers of the selected accents who could individually produce one of the selected accents precisely because they were native speakers of that particular accent. This is the most common research technique used in speech evaluation research focused on non-native English accents. Another assessment technique is the matched-guise technique that uses one bilingual or multilingual speaker who produces all language varieties. The main benefit of the matched-guise technique is that aspects such as voice quality are eliminated, which may potentially influence responses to speakers. Study 2 was focused on selecting matched guises for non-native and native English accents by means of a clearly defined and tested research methodology. The reason why in non-native English speech evaluation research matched guises are not always applied is that it is assumed to be impossible to be both a native and a non-native speaker of English.

The second study has proven that it is possible to select matched guises for Dutch-accented English and standard British and American English. The successful implementation of a matched-guise in our studies might not be possible for all sorts of L1 and L2 English combinations, but study 2 has shown that it might be worth the effort to at least attempt to develop matched guises for future L1 and L2 English accentedness research.

The matched guises that were used in studies 3 and 4 had to be evaluated by listener groups in four different countries (The Netherlands, Germany, Singapore, Spain). In order to be able to do this effectively, a global survey software and online data collection company, Qualtrics, was hired to collect the majority of the listener data. To this end, Qualtrics sampled highly educated listeners who were L1 speakers of the main national language of either the Netherlands (Dutch), Germany (German), Spain (Spanish), and Singapore (Singapore English). There were no restrictions with respect to the regions the listeners in the different countries came from. All respondents were registered individually, and all the data was checked by the main researcher after each data collection round, which resulted in three data collection rounds that took about two weeks per data collection round. Approximately 35% of the collected data was excluded because respondents had provided nonsensical answers or only provided neutral (mid-scale) answers to rating questions. The excluded data were replaced by new data. Although the data collection process was at times tedious, the international data collection overall took place relatively quickly and cheaply, and made it possible to collect the data for studying the effect of language varieties on different international listener groups within a short period of time, which future research can benefit from.

The fact that this thesis focused on highly educated listeners with at least average (self-reported) English fluency skills might mean that the results cannot be generalized to listeners with lower educational backgrounds and lower fluency levels in English. Future research should incorporate varying listener groups, with different educational and English proficiency backgrounds. For instance, listeners from nations that due to British colonial history have been exposed to English for centuries and today still use it as an important L2 in their government, education, and even literature, such as Kenia, Nigeria, and India, might respond differently to L1 and L2 English accents compared to L1 English listener groups, and other L2 English listener groups. Another relevant approach is to include listener groups from the major language families in the world. Such accent evaluation studies will allow researchers to better grasp what the current state of affairs is for *speech understandability* and *speaker evaluations* for both L2 and L1 English speakers, whether

language background matters, and whether the observed differences between L1 and L2 English listeners in our studies might illustrate a general pattern, which may be an indication of the growing relevance of an LFE speaking community. In these studies, research should focus on studying associations with both familiar and unfamiliar accents. This approach was taken by Yook and Lindemann (2012). They compared listeners' responses to English accent varieties on the basis of which listeners had and which listeners had not been notified on the nationality/ethnicity of the speakers. This research strategy allowed the researchers to understand whether the responses to the tested English accent varieties were based on associations with specific English accents, or not, and a similar research approach might yield valuable insights in the future.

The responses to accents were based on responses to native speakers' accents in the verbal guise study (study 1), and representative matched guises produced and tested in two other studies (studies 3 and 4). The intent in this thesis was to understand how listeners would respond to accents that were deemed generally representative for the selected accents, the speech samples were not chosen on the basis of exact non-native phonetic or prosodic elements. There have been a number of studies that have outlined features typical of a Dutch English pronunciation (e.g. Collins & Mees, 2003; Gussenhoven & Broeders, 1997; van den Doel, 2006), and some of them have developed a hierarchy of pronunciation errors to be avoided by non-native English speakers in general and Dutch English speakers in particular (Jenkins & Seidlhofer, 2001; van den Doel, 2006).Future studies should design test items based on established pronunciation 'error' hierarchies (e.g. van den Doel, 2006) to assess and compare the effects of specific pronunciation 'errors' on both L1 and L2 English listener groups.

This thesis focused on accentedness and communication context, and has shown that responses to accents should be considered in the context in which they are heard. The speech samples used to test the effect of the accents were based on scripts that were written to represent the selected communication contexts in terms of content and structure. They were of high quality and had limited distractions for listeners in terms of, for instance, background noise in order to allow us to measure accentedness responses under very good conditions. Other conditions, which are common in casual speech, such as reduced word pronunciations, noise might impact responses to accents. For example, research has shown that for L2 listeners reduced pronunciation can negatively impact speech understanding (Ernestus, Dikmans & Giezenaar, 2017; Ernestus & Warren, 2011). Therefore, future L2 and L1 English accentedness studies should consider studying responses to accents under varied conditions, for instance,

comparing speech produced with reduced word pronunciation variants with clearly articulated speech.

Dutch-accented English evoked responses that were relatively similar to standard British and American English. This might be connected to the fact that English and Dutch are part of the same language family, the West Germanic languages. As a result, Dutch-accented English might, linguistically speaking, be relatively similar to L1 English accents, compared with L2 English accents produced by speakers from other language families. Research has shown that linguistic distance can be a predictor of general difficulty of learning to speak a language (Schepens, van der Slik & van Hout, 2017). Language learning becomes more difficult as the linguistic distance between the L1 and L2 increases. In the context of this thesis this might mean that it is relatively easy for Dutch speakers to learn to speak English, because of the general similarities between English and Dutch. As a result the responses to Dutch-accented English in the present thesis might reflect responses to an L2 English accent that is relatively similar to two tested L1 English accents, and thus does not evoke strong responses in listeners with a completely different L1 background. Future research could, therefore, incorporate more L2 English accents that are selected on the basis of their linguistic features – more or less related to English – in order to further validate this thesis's results.

7
REFERENCES
SAMENVATTING
ACKNOWLEDGMENTS
CURRICULUM VITAE

REFERENCES

- Abu-Rabia, S. (2004). Teachers' Role, Learners' Gender Differences, and FL Anxiety Among Seventh-Grade Students Studying English as a FL. Educational Psychology, 24(5), 711-721, DOI: 10.1080/0144341042000263006.
- Ahn H. (2014). Teachers' attitudes towards Korean English in South Korea. *World Englishes, 33,* 195-222.
- Alexander, R. J. (1999). *Caught in a global English trap, or liberated by a lingua franca? Unravelling some aims, claims and dilemmas of the English teaching profession.* In Claus Gnutzmann (ed.), Teaching and Learning English as a Global Language: Native and Non-native perspectives (pp 23-39). Tubingen: Stauffenburg-Verlag.
- Baker, C. (1992). Attitudes and language. *Multilingual Matters 83.* Multilingual Matters, Clevedon.
- Bayyurt Y. (2018). Issues of intelligibility in world Englishes and EIL contexts. *World Englishes, 37,* 407– 415. https://doi.org/10.1111/weng.12327 .
- Bayard, D. & Green, J. (2005). *Evaluating English accents worldwide.* Te Reo, 48, 21-28.
- Bayard, D., Weatherall, A., Gallois, C., & Pittam, J. (2001). Pax Americana? Accent attitudinal evaluations in New Zealand, Australia, and America. *Journal of Sociolinguistics 5,* 22–49.
- Bent T, & Bradlow A. (2003). The interlanguage speech intelligibility benefit. *Journal of Acoustical Society of America, 114,* 1600-10.
- Bernaisch T. (2012). Attitudes towards Englishes in Sri Lanka. *World Englishes, 31,* 279-291. doi:10.1111/j.1467-971X.2012.01753.x.
- Berns M. (2018). Intelligibility and pedagogy. *World Englishes, 37,* 416-420. https://doi.org/10.1111/weng.12328.
- Bokhorst-Heng W, Caleon I. (2009). The language attitudes of bilingual youth in multilingual Singapore. *Journal of Multilingual and Multicultural Development, 30(3),* 235–251.
- Bolton, K. & Kuteeva, M. (2012). English as an academic language at a Swedish university: parallel language use and the 'threat' of English. *Journal of Multilingual andMulticultural Development, 33(5),* 429-447. DOI: 10.1080/01434632.2012.670241.
- Bolton K., & De Costa P. (2018). World Englishes and Second Language Acquisition: Introduction. *World Englishes, 37,* 2-4. DOI: 10.1111/weng.12298.
- BON (2017). Verengelsing hoger onderwijs: steun onze rechtszaak. Beter Onderwijs Nederland, 18 May 2018, retrieved on 22 April 2019, from https://www.beteronderwijsnederland.nl/nieuws/2018/05/rechtszaak.
- Bongaerts, T., van Summeren, C., Planken, B., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition, 19(4),* 447-465. doi:10.1017/S0272263197004026.

- Bouma, K., (2016). Meer dan de helft van de studies volledig in het Engels. Volkskrant, 26 August 2016, retrieved on 22 April 2019, from https://www.volkskrant.nl/binnenland/meer-dan-de-helft-van-de-studies-volledig-in-het-engels~a4364526.
- Bouma, K., (2018). Sommige studies zijn zo in trek dat universiteiten een stop op buitenlandse studenten willen .Volkskrant, 8 March 2018, retrieved on 22 April 2019, from https://www.volkskrant.nl/nieuws-achtergrond/sommige-studies-zijn-zo-in-trek-dat-universiteiten-een-stop-op-buitenlandse-studenten-willen~b76321b2.
- British Council, (2017). IELTS International English language test system. Retrieved 30 November 2018: http://takeielts.britishcouncil.org.
- British Library (2019). Minority Ethnic English. British library. [cited 7 July 2019]. Available from https://www.bl.uk/british-accents-and-dialects/articles/minority-ethnic-english.
- Bronkhorst, X. (2015). DUB-panel: Engels als onderwijstaal? Meer voor- dan nadelen. University Utrecht, 6 May 2015, retrieved on 22 April 2019, from https://www.dub.uu.nl/nl/achtergrond/dub-panel-engels-als-onderwijstaal-meer-voor-dan-nadelen.
- Brown, R. (1965). *Social Psychology.* New York: Macmillan.
- Brown A. (1999). *Singapore English in a nutshell: An alphabetical description of its features.* Federal Publications.
- Burgoon JK, Burgoon M. (2001). *Expectancy theories.* In W.P. Robinson, & H. Giles (Eds.), The new handbook of language and social psychology (2nd ed., pp 79-102). Sussex: Wiley; Cambridge English. N.d. Retrieved 14 October, 2019: https://www.cambridgeenglish.org/why-choose-us.
- Burgoon, J.K. & Le Poire, B.A. (1999). Nonverbal cues and interpersonal judgments: Participant and observer perceptions of intimacy, dominance, composure, and formality. *Communication Monographs, 66:2*, 105-124, DOI: 10.1080/03637759909376467.
- Burgoon J.K., Stacks D.W. & Woodall W.G. (1979). A communicative model of violations of distancing expectations. *Western Journal of Speech Communication, 43,* 153-167.
- Canagarajah S. (2007). Lingua Franca English, Multilingual Communities, and Language Acquisition. *The Modern Language Journal, 91,* 923-939. doi:10.1111/j.1540-4781.2007.00678.x.
- Caraker R. (2016). Spain and the context of English language education. *Research Bulletin, 92,* 23-35.
- Cargile, A. (1997). Attitudes toward Chinese-accented speech: An investigation in two contexts. *Journal of Language and Social Psychology, 16,* 434–443.
- Cargile, A. C. & Howard G. (1997). Understanding language attitudes: Exploring listener affect and identity. *Language and Communication 17,* 195-217.
- Cargile, A., & Giles, H. (1998). Language attitudes toward varieties of English: An American Japanese context. *Journal of Applied Communication Research, 26,* 338-356.

- Cavallaro, F., Ng, B. & Seilhamer, M. (2014). Singapore Colloquial English: Issues of prestige and identity. *World Englishes 33(3),* 378–397. DOI:10.1111/weng.12096.
- Central Bureau for Statistics, Netherlands (2019). Internationale handel. CBS Statline [cited 7 July 2019]. Available from https://opendata.cbs.nl/statline/#/CBS/nl/dataset/82616NED/table?ts=1556200130662.
- Chiba R , Matsuura H. & Yamamoto A. (1995). Japanese attitudes toward English accents. *World Englishes, 14,* 77-86. doi:10.1111/j.1467-971X.1995.tb00341.x.
- Cohen, Jacob (1988). *Statistical Power Analysis for the Behavioral Science,* 2nd edn. New York: Academic Press.
- Collins, B., Mees, I., (2003). *The Phonetics of English and Dutch.* Fifth Revised Edition. Brill, Leiden/Boston.
- Corder SP. (1967). The Significance of Learners' Errors. *International Review of Applied Linguistics in Language Teaching, 5,* 161-170. DOI: http://dx.doi.org/10.1515/iral.1967.5.1-4.161.
- Council of Europe, (2017). Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Strasbourg: Language Policy Division, Council of Europe. Retrieved 30 November 2018: http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf.
- Coupland, N., & Bishop, H. (2007). Ideologised values for British Accents. *Journal of Sociolinguistics, 11,* 74-93.
- Coupland, N. & T. Kristiansen. (2011). *SLICE. Critical perspectives on language (de)standardisation.* In T. Kristiansen, T., Coupland, N. (eds.) Standard Languages and Language Standards in a Changing Europe. Novus, Oslo, pp. 11-35.
- Crystal D. (2009). *English as a Global Language.* 2nd Edition. Cambridge: Cambridge University Press.
- Dalton-Puffer, C., Kaltenboeck, G., Smit, U., 1997. Learner Attitudes and L2 Pronunciation in Austria. *World Englishes 16(1),* 115-128. DOI:10.1111/1467-971X.00052.
- De Groot, E. (2008) *English annual reports in Europe: A study on the identification and reception of genre characteristics in multimodal annual reports originating in the Netherlands and in the United Kingdom.* PhD, Utrecht.
- Department of Statistics, Singapore (2018). Education, Language Spoken and Literacy. [cited 7 July 2019]. Available from https://www.singstat.gov.sg/find-data/search-by-theme/population/education-language-spoken-and-literacy/latest-data.
- Derwing TM, Munro MJ. (1997). Accent, intelligibility and comprehensibility: evidence from four L1s. *Studies in second language acquisition,* 19, 1-16.
- Derwing, T., & Munro, M. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching, 42(4),* 476-490.

- De Volkskrant (2006) Indiërs openen callcenters in Belfast. 10 August. Retrieved 6 April 2011 fromhttps://www.volkskrant.nl/nieuws-achtergrond/indiers-openen-callcenter-in-belfast~b2cf7019.
- de Vos, J. (2019) Thesis defense Johanna de Vos (Donders series 379). Retrieved 14 October, 2019: https://www.ru.nl/donders/graduate-school/after-your-phd/donders-thesis-series/virtuele-map/thesis-defense-johanna-de-vos-donders-series-379.
- Edwards, A., (2016). *English in the Netherlands: Functions, forms and attitudes.* John Benjamins Publishing Company, Amsterdam. DOI: 10.1075/veaw.g56.
- El-Dash, L., Busnardo, J., (2001). Brazilian attitudes toward English: Dimensions of status and solidarity. *International Journal of Applied Linguistics 11(1),* 57-74. Retrieved 30 November 2018: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1473-4192.00004.
- EF (2017). EP English Proficiency Index. Retrieved 30 November 2018: https://www.ef.nl/epi.
- EF (2018). EF English Proficiency Index. EF, 2018, retrieved on 22 April 2019, from https://www.ef.com/wwen/epi.
- Ellis R. (1994). *The study of second language acquisition.* London: Oxford University Press; 1994.
- Ernestus, M. Dikmans M., & Giezenaar, G (2017). Advanced second language learners experience difficulties processing reduced word pronunciation variants. *Dutch Journal of Applied Linguistics 6 (1),* 1-20. doi:10.1075/dujal.6.1.01ern.
- Ernestus, M. & Warner, N (2011). An introduction to reduced pronunciation variants. *Journal of Phonetics 39,* 253-260. doi:10.1016/S0095-4470(11)00055-6.
- ETS. (n.d.). Retrieved 14 October 2019: https://www.ets.org/about.
- European Commission (2006). Special eurobarometer: Europeans and their languages. Retrieved 26 September 2019 from http://ec.europa.eu/public_opinion/archives/ebs/ebs_243_en.pdf.
- Fayer, JM., & Krasinski, E. (1987) Native and non-native judgments of intelligibility and irritation. *Language Learning 37,* 313–325.
- Federal Foreign Office, Germany, (2019). Germany and the Netherlands: bilateral relations. [cited 7 July 2019]Available from https://www.auswaertiges-amt.de/en/aussenpolitik/laenderinformationen/niederlande-node/germany-netherlands-bilateral/227968#content_2.
- Firth A, Wagner J.(1997). On discourse, communication, and (some) fundamental concepts in SLA research. *Modern Language Journal, 81,* 285-300.
- Flowerdew, John (ed.) (1994). *Academic Listening.* Cambridge University Press: Cambridge.
- Friedmann, A. & Rusou, D. (2015). Critical period for first language: the crucial role of language input during the first year of life. *Current Opinion in Neurobiology, 35:* 27-34,https://doi.org/10.1016/j.conb.2015.06.003.

- Garrett, P., (2010). *Attitudes to Language.* Key topics in sociolinguistics. Cambridge University Press, Cambridge.
- Gass SM, & Selinker L. (1994). *Second language acquisition: an introductory course.* 3rd ed. New York: Routledge.
- Gerritsen, M., Korzilius, H., Van Meurs, F., & Gijsbers, I. (2000). English in Dutch commercials: Not understood and not appreciated. *Journal of Advertising Research, 40,* 17-36.
- Gerritsen, M. & Nickerson, C. (2004). *Fact or fallacy? English as a lingua franca in the European business context.* In Christopher N. Candlin, and Maurizio Gotti (eds.), Intercultural Aspects of Specialized Discourse (pp. 105-125). Berlin: Lang.
- Gerritsen, M., Nickerson, C., van den Brandt, C., Crijns, R., Dominguez, N., & van Meurs, F., et al. (2007). *English in print advertising in Germany, Spain and the Netherlands: Frequency of occurrence, comprehensibility and the effect on corporate image.* In Giuliana Garzone and Cornelia Ilie (eds.), The Role of English in Institutional and Business Settings. (pp. 79-98). Berlin: Peter Lang.
- Gerritsen M, Nickerson C, Van Hooft A, van Meurs F, Korzilius H, Nederstigt U. et al. (2010). English in Product Advertisements in Non-English-Speaking Countries in Western Europe: Product Image and Comprehension of the Text. *Journal of Global Marketing, 23,* 1-17.
- Gerritsen, M., Van Meurs, F., Planken, B., Korzilius, H., (2016). A reconsideration of the status of English in the Netherlands within the Kachruvian Three Circles model. *World Englishes 35,* 457-474. http://dx.doi.org/10.1111/weng.12206.
- Gijsberts M, Lubbers M, Fleischmann F, Maliepaard M., & Schmeets H. SCP (2016). [cited 7 July 2019]. Available from https://www.scp.nl/Publicaties/Alle_publicaties/Publicaties_2016/Nieuwe_Spaanse_migranten_in_Nederland.
- Giles, H., (1970). Evaluative reactions to accents. *Educational Review 22 (3),* 211-227. http://dx.doi.org/10.1080/0013191700220301.
- Giles, H., & Powesland, P.F. (1975). *Speech Style and Social Evaluation.* London: Academic Press.
- Greven K. (2016). Spanje krijgt dit jaar recordaantal toeristen. NRC. [cited 7 July 2019]. Available from https://www.nrc.nl/nieuws/2016/07/11/met-zn-68-miljoen-op-vakantie-naar-spanje-3203669-a1510832.
- Grondelaers, S., Kristiansen, T., (2013). *On the need to access deep evaluations when searching for the motor of standard language change.* In: Grondelaers, S., Kristiansen, T., (Eds.), Language (De)standardisation in Late Modern Europe. Experimental Studies, pp. 9-52.
- Grondelaers S, van Hout R. (2015). How (in)coherent can standard languages be? A perceptual perspective on co-variation. *Lingua, 172.* DOI: 10.1016/j.lingua.2015.10.012.
- Grondelaers, S., Van Gent, P., & van Hout, R., (2015). *Is Moroccan-flavoured standard Dutch standard or not? On the use of perceptual criteria to determine the limits of standard*

*languages. I*n: Prikhodkine, A., Preston, D.R. (Eds.), Responses to Language Varieties: Variability, Processes and Outcomes. John Benjamins, pp. 191-218. http://dx.doi.org/10.1075/impact.39.

- Gussenhoven, C. & Broeders, T. (1997). *English Pronunciation for Student Teachers,* 2nd edn. Groningen: Wolters-Noordhof.
- Guy, G., and Cutler, C., (2011). Speech style and authenticity: Quantitative evidence for the performance of identity. *Language Variation and Change, 23.* DOI: 10.1017/S0954394510000232.
- He D, Miller L. (2011). English teacher preference: the case of China's non-English-major students. *World Englishes, 30:* 428-443. doi:10.1111/j.1467-971X.2011.01716.x.
- He, D., & Zhang, Q., (2010). Native speaker norms and China English: From the perspective of learners and teachers in China. *TESOL Quarterly, 44(4),* 769-789.
- Hellekjær, G.O. (2010). Lecture comprehension in English-medium higher education. *Hermes-Journal of Language and Communication Studies, 45,* 11-34.
- Hendriks B, van Meurs F, & Hogervorst N. (2016) Effects of degree of accentedness in lecturers' Dutch-English pronunciation on Dutch students' attitudes and perceptions of comprehensibility. *Dutch Journal of Applied Linguistics, 5(1),* 1-17. DOI: https://doi.org/10.1075/dujal.5.1.01hen.
- Hendriks B, van Meurs F, & de Groot E. (2017). The effects of degrees of Dutch accentedness in ELF and in French, German and Spanish. *International Journal of Applied Linguistics; 27(1),* 44-66. DOI: 10.1111/ijal.12101.
- Hendriks B, van Meurs F, & Reimer AK. (2018). The evaluation of lecturers' nonnative-accented English: Dutch and German students' evaluations of different degrees of Dutch-accented and German-accented English of lecturers in higher education. *Journal of English for Academic Purposes, (34),* 28-45, DOI: https://doi.org/10.1016/j.jeap.2018.03.001.
- Huygen, M. (2017). Uhm...how do you say that in English? NRC Weekend. retrieved 22 April 2019: https://www.nrc.nl/nieuws/2017/10/21/uhm-how-do-you-say-that-in-english-13593132-a1578131.
- Jenkins, J. & Seidlehofer, B. (2001). Bringing Europe's lingua franca into the classroom. Retrieved 14 October 2019: https://www.theguardian.com/education/2001/apr/19/languages.highereducation1.
- Jenkins, J. (2006). Points of view and blind spots: ELF and SLA. *International Journal of Applied Linguistics 16,* 138–162.
- Kankaanranta A, & Planken B.(2010). Belf Competence as Business Knowledge of Internationally Operating Business Professionals. *Journal of Business Communication; 47(4),* 380-407. ISSN 0021-9436 (printed). DOI: 10.1177/0021943610377301.
- Kachru, Braj B. (1983/2). *The Other Tongue: English across Cultures.* Oxford: Pergamon.

- Kachru, Yamuna (2008). Cultures contexts, and interpretability. *World Englishes 27,* 309-318.
- Kachru, Y., & Smith, L. (2008). *Cultures, Contexts, and World Englishes.* New York: Routledge.
- Kalin, R., Rayko, DS., & Love, N. (1980). *The perception and evaluation of job candidates with four different ethnic accents'.* In Howard Giles, W. Peter Robinson, and Phillip Smith (eds.), Language: Social Psychological Perspectives (pp. 197–202). Oxford: Pergamon.
- Koet, Ton (2007). Polder English in Dutch ears: Empirical studies on the evaluation of the pronunciation of English as a Foreign Language. Retrieved 6 April 2011 from http://dare.uva.nl/document/51192.
- Koster, CJ., & Koet, T. (1993). The evaluation of Accent in the English of Dutchman. *Language Learning 43,* 69-92.
- Kristiansen G, Zenner E, & Geeraerts D. (2018). English as a Lingua Franca in Europe: the Identification of L1 and L2 accents: A Multifactorial Analysis of Pan-European Experimental Data. *Annual Review of Cognitive Linguistics, 16(2),* 494-518. https://doi.org/10.1075/rcl.00019.kri.
- Lambert, W., Hodgson, R., Gardner, R., Fillenbaum, S., (1960). Evaluational reactions to spoken languages. *The Journal of Abnormal and Social Psychology 60(1),* 44-51.
- Lee P. (2016). English most common home language in Singapore, bilingualism also up: Government survey. The Strait Times [cited 7 July 2019]. Available from https://www.straitstimes.com/singapore/english-most-common-home-language-in-singapore-bilingualism-also-up-government-survey.
- Lenneberg, E.H. (1967). *Biological foundations of language.* New York: Wiley and Sons Inc.
- Lev-Ari, S., & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology, 46(6),* 1093-1096.
- Lightbown PM, Spada N. (1999). *How languages are learned.* Oxford: Oxford University Press.
- Lindemann, S. (2002). Listening with an attitude: A model of native-speaker comprehension of non-native speakers in the United States. *Language in Society, 31,* 419-441.
- Lindemann, S. (2003). Koreans, Chinese, or Indians? Attitudes and ideologies about non-native English speakers in the United States. *Journal of Sociolinguistics, 7(3),* 348-364. DOI: https://doi.org/10.1111/1467-9481.00228.
- Lippi-Green, R., (1997). *English with an accent.* In: Language, Ideology, and Discrimination in the United States. Routledge, London/New York.
- Lizzini, O., Martijn , M., Munk , R., & De Regt, H. (2017). Academisch onderwijs kan niet zonder het Engels. Volkskrant, retrieved 22 April 2019, from https://www.volkskrant.nl/columns-opinie/academisch-onderwijs-kan-niet-zonder-het-engels~bcc9ca31.

- Mai, R., & Hoffmann, S., (2014). Accents in business communication: an integrative model and propositions for future research. *Journal of Consumer Psychology, 24(1),* 137–158. DOI:10.1016/j.jcps.2013.09.004.
- MacFarlane, A., & Stuart-Smith, J., (2012).'One of them sounds sort of Glasgow Uni-ish'. Social judgements and fine phonetic variation in Glasgow. *Lingua 122 (7),* 764-778. DOI:10.1016/j.lingua.2012.01.007.
- Major, RC., Fitzmaurice, SM., Bunta, F, & Balasubramanian, CC. (2005) Testing the effects of regional, ethnic, and international dialects of English on listening comprehension. *Language Learning 55,* 37-69.
- Matsuura, H., Chiba, R., & Yamamoto, A. (1994). Japanese college students attitudes towards non-native varieties of English. In D. Graddol, & J. Swann (Eds.), Evaluating Language (pp.52-61). Clevedon: BAAL/Multilingual Matters.
- McDonald C, & McRae S. (2010). A pre-trial collection and investigation of what perceptions and attitudes of Konglish exist among foreign and Korean English language teachers in terms of English education in Korea. *Asian EFL, 12(1),* 134-164.
- McKenzie RM. (2008). The role of variety recognition in Japanese university students' attitudes towards English speech varieties. *Journal of Multilingual and Multicultural Development, 29(2),* 139-153. doi:10.1111/j.1473-4192.2008.00179.x.
- Menzel, P., (2017). Robosapiens. Cambridge MA: The MIT Press. Retrieved 30 November 2018: http://robosapiens.mit.edu/gecko.htm.
- Meyerhoff, M. (2011). *Introducing Sociolinguistics.* Routledge, London.
- Milroy, J. (1992). *Linguistic Variation and Change.* Blackwell, Oxford/Cambridge, MA.
- Munro, MJ., & Derwing TM. (1995a) Foreign accent, comprehensibility and intelligibility in the speech of second language learners. *Language Learning, 45,* 73-97.
- Munro, M J., & Derwing TM. (1995b) Processing time, accent and comprehensibility in the perception of native and foreign accented speech. *Language and Speech, 38,* 289-306.
- Munro, M J., Derwing, TM., & Morton, SL. (2006) The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition, 28,* 111-131.
- Nelson, C.L. (2011). *Intelligibility in World Englishes: theory and application.* New York & London: Routledge.
- Nickerson, C. (2005). English as a lingua franca in international business contexts. *English for Specific Purposes, 24(4),* 367-380.
- Orikasa, M., (2016). The intelligibility of varieties of English in Japan. *World Englishes 35(3),* 355–371. DOI:10.1111/weng.12209.
- Penning de Vries BW., Cucchiarini C, Strik H, & van Hout R. (2019). Spoken grammar practice in CALL: The effect of corrective feedback and education level in adult L2 learning. *Language Teaching Research.* DOI: https://doi.org/10.1177/1362168818819027.

- Pennycook, A. (1998) *English and the Discourses of Colonialism.* London: Routledge.
- Phillipson, R. (1992). *Linguistic Imperialism.* Oxford: Oxford University Press.
- Pihko, M. K. (1997). *'His English sounded strange': The intelligibility of native and non-native English pronunciation to Finnish learners of English.* Doctoral dissertation University of Jyvaskyla, Finland, Dissertation Abstracts International 60, 9C.
- Preston, D.R., (1989). *Standard English spoken here: The geographical loci of linguistic norms.* In Ammon, Ulrich (ed.), Status and function of languages and language varieties. Walter de Gruyter, Berlin/New York, pp. 324–54. ISBN: 9783110112993.
- Purnell, T., Idsardi, W., & Baugh, J., (1999). Perceptual and Phonetic Experiments on American English Dialect Identification. *Journal of Language and Social Psychology 18 (1),* 10-30. DOI: 10.1177/0261927x99018001002.
- Qualtrics. Survey Software. Qualtrics [cited 7 July 2019]. Available from https://www.qualtrics.com/research-core/survey-software.
- Richards JC. (1974). *Error Analysis: Perspectives on second language acquisition.* London: Longman.
- Rijksdienst Voor Ondernemend Nederland (RVO). (2019a). Duitsland. [cited 7 July 2019]. Available from https://www.rvo.nl/onderwerpen/internationaal-ondernemen/landen-overzicht/duitsland.
- Rijksdienst Voor Ondernemend Nederland (RVO). (2019b). Spanje. [cited 7 July 2019]. Available from https://www.rvo.nl/onderwerpen/internationaal-ondernemen/landen-overzicht/spanje/handel-nederland-spanje.
- Rindal, U., Piercy, C., (2013). Being 'neutral'? English pronunciation among Norwegian learners. World Englishes 32(2), 211–229. DOI:10.1111/weng.12020.
- Rivers DJ. (2011). Intercultural processes in accented English. *World Englishes, 30,* 375-391. doi:10.1111/j.1467-971X.2011.01707.x.
- Rosen, M. (2010). Uhm... how do you say that in English? Retrieved 14 October 2019: https://www.bbc.com/news/magazine-12017753.
- Rubin, DL. (1992) Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. *Research in Higher Education 33,* 511-531.
- Ryan, E. B., & Bulik, C. M. (1982). Evaluations of middle class and lower class speakers of standard American and German-accented English. *Journal of Language and Social Psychology, 1(1),* 51-61.
- Ryan, E. B., & Sebastian, R. J. (1980). The effects of speech style and social class background on social judgements of speakers. *British Journal of Social and Clinical Psychology, 19,* 229-33.
- Ryan, E. B., & Giles, Howard (1992). *Attitudes towards Language Variation.* London: Edward Arnold.

- Sasayama S. (2013). Japanese college students' attitudes towards Japan English and American English. *Journal of Multilingual and Multicultural Development, 34(3),* 264-278, DOI:10.1080/01434632.2013.767341.
- Schepens, J., van der Slik, F., & van Hout, R. (2013). *The effect of linguistic distance across Indo-European mother tongues on learning Dutch as a second language.* Approaches to Measuring Linguistic Differences (pp. 199-230).
- Schüppert, A., Hilton, N., & Gooskens, C., (2015). Swedish is beautiful, Danish is ugly? Investigating the link between language attitudes and intelligibility. *Linguistics, 53(2),* 275-304. DOI: 10.1515/ling-2015-0003.
- Seidlhofer B, Breiteneder A, & Pitzl ML. (2006) English as a lingua franca in Europe: Challenges for applied linguistics. *Annual Review of Applied Linguistics, 26,* 3-34.
- Selinker L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching, 10,* 209-241. DOI: http://dx.doi.org/10.1515/iral.1972.10.1-4.209.
- Smakman, D., (2006). *Standard Dutch in the Netherlands. A sociolinguistic and phonetic description.* Dissertation University of Nijmegen. LOT, Utrecht.
- Smakman, D., (2012). The Definition of the Standard Language. A Survey in Seven Countries. *International Journal of the Sociology of Language 218,* 25-58. DOI: 10.1515/ijsl-2012-0058.
- Smakman, D. & Barasa, S.N., (2017). *Defining 'Standard': Towards a Cross-cultural Definition of the Language Norm.* In: Tieken-Boon van Ostade, I., Percy, C. (Eds.), Prescription and Tradition in Language: Establishing Standards across Time and Space. Multilingual Matters, Bristol, pp. 23-38.
- Smith, LE., & Bisazza, JA. (1982) The comprehensibility of three varieties of English for college students in seven countries. *Language Learning, 32,* 259-269.
- Smith, LE. (1992). *Spread of English and issues of intelligibility.* In Braj B. Kachru (ed.), The Other Tongue: English across culture, 2nd edn (pp. 75-90). Urbana, IL: University of Illinois Press.
- Smith, LE., & Nelson, CL. (2006). *World Englishes and issues of intelligibility.* In Braj B. Kachru, Yamuna Kachru, and Cecil L. Nelson (eds.), The Handbook of World Englishes (pp. 428-445). Malden, MA: Blackwell Publishing.
- Sociaal en Cultureel Planbureau (2005). *De Sociale Staat van Nederland.* Den Haag: Sociaal en Cultureel Planbureau.
- Tokumoto, M. & Shibata, M., (2011). Asian varieties of English: Attitudes towards pronunciation. *World Englishes 30(3),* 392–408. DOI:10.1111/j.1467-971X.2011.01710.x.
- Trudgill, P. (2001). *Sociolinguistic Variation and Change.* Edinburgh University Press, Edinburgh.
- Trudgill P, & Hannah J. (2008). *International English: A Guide to Varieties of Standard English*. Routledge. 5th Ed. London and New York: Routledge.

- Tsalikis, J., deShields Jr, O. W., & La Tour, M. S. (1991). The role of accent on the credibility and effectiveness of the salesperson. *Journal of Personal Selling and Sales Management, 11(19),* 31-41.
- Van den Doel, R., (2006). *How Friendly are the natives? An evaluation of native-speaker judgments of foreign-accented British and American English.* LOT Publications, Utrecht.
- Van der Haagen, M. (1998). Caught between Norms: The English Pronunciation of Dutch Learners. The Hague: Holland Academic Graphics.
- Van Gaal, J. (2018). Engels moet voertaal TU/e worden. Retrieved 22 April 2019, from https://www.cursor.tue.nl/nieuws/2018/januari/week-4/engels-moet-voertaal-tue-worden.
- Van Heest, H. (2018). Minister overweegt de wet te wijzigen over verengelsing. Scienceguide, 4 April 2018, retrieved 22 April 2019, from https://www.scienceguide.nl/2018/04/minister-overweegt-de-wet-te-wijzigen-over-verengelsing.
- Van Onna, B. & Jansen, C. (2006). Nederland talenland? Over de beheersing van Engels, Duits, Frans en Nederlands in Nederlandse organisaties. Retrieved 6 April 2011 from http://cjpublications.nl/2006_VanOnna_Jansen-Nederland%20talenland.pdf.
- Van Oostendorp, M. (2002). *Steenkolen Engels.* Amsterdam: Uitgeverij Veen.
- Varonis, EM., & Gass, S., (1982) The comprehensibility of non-native speech. *Studies in Second Language Acquisition, 4,* 114-136.
- Vermeulen, R., & Kellerman, E. (1998). *Causation in narrative: The role of language background and proficiency in two episodes of 'the frog story'.* In Dorte Albrechtsen, Birgit Henricksen, Inger M. Mees, and Poulsen, Erik (eds.), Perspectives on Foreign and Second Language Pedagogy (pp. 161–176). Odense: Odense University Press.
- VNO-NCW: The Confederation of Netherlands Industry and Employers. (2018). Samenwerking met Singapore kans in Brexit-problematiek. [cited 7 July 2019]. Available from https://www.vno-ncw.nl/nieuws/samenwerking-met-singapore-kans-brexit-problematiek.
- Wang, H., (2007). *English as a Lingua Franca: Mutual Intelligibility and American Speakers of English.* Utrecht: Netherlands Graduate School of Linguistics.
- Wang, H, & Van Heuven, VJ. (2007). *Quantifying the interlanguage speech intelligibility benefit.* In William, J. Barry, and Jürgen Trouvain (eds.), Proceedings of the 16th International Congress of Phonetic Sciences. Saarbrücken (pp. 1729-1732). Saarbrucken: Universität des Saarlandes.
- Wilcox, George K. (1978). The effect of accent on listening comprehension: A Singapore study. *English Language Teaching Journal 32,* 118-127.
- Wolfram, W., Schilling-Estes, N. (2006). *American English: Dialects and variation* (2nd ed. ed., Language in society, 25). Blackwell Pub, Malden, MA.

- Yook, C., & Lindemann, S. (2013). The role of speaker identification in Korean university students' attitudes towards five varieties of English. *Journal of Multilingual and Multicultural Development, 3(34),* 279-296.
- Yorkston, K., Strand, E., & Kennedy, M. (1996). Comprehensibility of Dysarthric Speech: Implications for Assessment and Treatment Planning. *American Journal of Speech-Language Pathology, 5,* 55-66.
- Zahn, C. J., & Hopper, R. (1985). Measuring language attitudes: The Speech Evaluation Instrument. *Journal of Language and Social Psychology, 4,* 113-123.

SAMENVATTING

Engels is een wereldtaal die veelvuldig als lingua franca wordt gebruikt wanneer niet-moedertaalsprekers met elkaar communiceren. Het aantal leerders van het Engels groeit daardoor nog steeds en Engels is in veel samenlevingen de belangrijkste vreemde taal geworden. Zo ook in Nederland waar het de voornaamste vreemde taal is in de communicatie van overheid, media, wetenschap, onderwijs en bedrijfsleven. Er bestaat in Nederland ook discussie over de status van het Engels binnen deze domeinen en welke variëteit van het Engels geleerd en gebruikt zou moeten worden om zo goed mogelijk in het Engels te kunnen communiceren met moedertaalsprekers, met andere Nederlanders en andere niet-moedertaalsprekers van het Engels. Nederlanders hebben allerlei opvattingen over de gewenste dan wel noodzakelijke kwaliteit van hun eigen Engels, maar die zijn lang niet altijd gebaseerd op taalkundige inzichten.

Traditioneel gezien wordt er in tweedetaalverwervingonderzoek vanuit gegaan dat moedertaalsprekers 'eigenaar' en gebruiker zijn van een taal, terwijl niet-moedertaalsprekers een taal leren. Tweedetaalverwervingsonderzoek houdt zich onder andere bezig met het bestuderen van de taalproductie van niet-moedertaalsprekers en hoe deze afwijkt van de taalproductie van moedertaalsprekers. De meest succesvolle taalstudent wordt in algemene zin gedefinieerd als diegene wiens taalproductie zo min mogelijk afwijkt van de taalproductie van moedertaalsprekers. Een aantal taalkundigen denkt dat de huidige positie van Engels als wereldtaal en belangrijke lingua franca heeft geleid tot het ontstaan van een internationale lingua franca Engels (LFE) taalgemeenschap waarbinnen het traditionele tweedetaalverwervingsperspectief niet relevant is. Deze LFE taalgemeenschap bestaat uit alle niet-moedertaalsprekers van het Engels die taalvaardig genoeg zijn om in het Engels te kunnen communiceren met andere niet-moedertaalsprekers van het Engels, en met moedertaalsprekers van het Engels. De LFE taalgemeenschap kenmerkt zich door het vermogen om variëteiten van het Engels te begrijpen en een hoge tolerantie ten aanzien van de verschillende variëteiten van het Engels. Om erachter te komen hoe Nederlanders die Engels spreken, begrepen en gepercipieerd worden door moedertaalsprekers van het Engels, door andere Nederlanders, en door andere niet-moedertaalsprekers van het Engels zijn vier onderzoeken uitgevoerd. De vier onderzoeken in deze dissertatie zijn gericht op het effect van een Nederlands accent in het Engels. Hoe positief of negatief wordt een Nederlands accent beoordeeld?

Hoofdstuk 2 beschrijft een experiment waarin vergeleken werd hoe hoogopgeleide Britse luisteraars reageerden op twee gradaties van een Nederlands-Engels accent (*licht* en *gematigd*) en een standaard Brits-Engels accent. De spraakfragmenten waren volgens de verbal-guise methode geproduceerd, wat wil zeggen dat meerdere sprekers van deze drie accenten waren geselecteerd om de betreffende accenten zo natuurlijk mogelijk te produceren. De luisteraars (N=144) moesten de inleiding van een telefonisch pitch beluisteren, de sprekers evalueren en begripsvragen beantwoorden. De spraakevaluaties waren gericht op de sociale status (*status*) van de sprekers en het hebben van een prettige persoonlijkheid (*affectiviteit*). De begripsvragen (*spraakbegrip*) waren gericht op de volgende drie niveaus : (1) *verstaanbaarheid,* ofwel het verstaan van woorden en zinnen; (2) *begrip,* ofwel het begrijpen van de woorden en zinnen in letterlijke en contextuele zin; en (3) *interpretatie*, het kunnen interpreteren van het doel van de boodschap en de intenties van de spreker. De groet 'how are you' is bijvoorbeeld correct verstaan (*verstaanbaarheid*) wanneer een luisteraar ieder woord correct kan transcriberen, goed begrepen (*begrip*) wanneer een luisteraar begrijpt wat de letterlijke betekenis is van de woorden en de complete groet en dat het een veelvoorkomende manier van gedag zeggen is. Het goed interpreteren (*interpretatie*) van een groet als 'how are you' is het onderkennen dat dit in Angelsaksische culturen een sociale groet is die niet als doel heeft om een gesprek aan te gaan over allerlei persoonlijke kwesties. Dat wordt door niet-moedertaalsprekers niet altijd begrijpen.

De resultaten lieten zien dat door Britse luisteraars aan de sprekers met een standaard Brits-Engels accent een hogere status werd toegekend dan aan de sprekers met het lichte en het gematigde Nederlands-Engelse accent. Daarnaast werden de sprekers met een standaard Brits-Engels en een licht Nederlands-Engels accent als prettiger en vriendelijker ervaren dan de sprekers met een gematigd Nederlands-Engels accent. Hoewel de Britse luisteraars het standaard Brits-Engelse accent verstaanbaarder en begrijpelijker vonden dan beide Nederlands-Engelse accenten, konden ze de boodschap inhoudelijk in alle drie de accenten even goed interpreteren.

Op basis van het eerste experiment is een *matched-guise* methode voor accenten in het Engels ontwikkeld en getoetst in het tweede experiment (Hoofdstuk 3). Een matched-guise is spraak van één spreker die verschillende accenten kan produceren omdat hij of zij meestal twee- of meertalig is. Een groot voordeel van matched-guise spraakfragmenten is dat luisteraars uitsluitend reageren op de verschillende accenten en niet op verschillende stemmen, omdat dezelfde spreker de verschillende accenten produceert. Het vinden en het selecteren van matched-guise sprekers voor onder-

zoek gericht op het vergelijken van reacties op moedertaal en niet-moedertaal Engelse accenten is niet eenvoudig omdat iemand niet tegelijk een moedertaalspreker en een niet-moedertaalspreker kan zijn van verschillende accenten. In het tweede experiment is onderzocht of we sprekers konden vinden die toch drie verschillende accenten konden produceren: een typisch Nederlands-Engels accent, een standaard Brits-Engels accent en een standaard Amerikaans Engels accent. Als dat zou lukken, dan zouden we deze methode kunnen inzetten in ons verdere onderzoek naar het effect van verschillende Engelse accenten.

Er zijn matched-guise spraakfragmenten opgenomen voor een 'typisch' Nederlands-Engels accent, een standaard Brits-Engels accent en een standaard Amerikaans Engels accent. Vier sprekers die volgens taalexperts de drie accenten konden produceren als een moedertaalspreker werden voor dit experiment geselecteerd. De opnames van de vier matched-guise sprekers zijn vervolgens door 'taalkundig naïeve' luisteraars geëvalueerd, dat wil zeggen luisteraars die geen taalkundige achtergrond hadden. Veertig moedertaalsprekers van het Nederlanders evalueerden het 'typisch' Nederlands-Engelse accent, 40 moedertaalsprekers van Brits-Engels in het Verenigd Koninkrijk evalueerden het standaard Brits-Engelse accent en 40 moedertaalsprekers van Amerikaans Engels in de Verenigde Staten evalueerden het standaard Amerikaans-Engelse accent. De spraakfragmenten van de vier matched-guise sprekers werden beoordeeld op de mate waarin een spreker werd beschouwd als een moedertaalspreker ('native') en de mate waarin de spreker werd beschouwd als een spreker met een standaard ('standard' voor de standaard Brits/Amerikaans-Engelse accenten) of typisch ('typical' voor het Nederlands-Engelse accent) accent.

De resultaten lieten zien dat het oordeel van de taalexperts niet altijd overeenkwam met die van de luisteraars. In tegenstelling tot de taalexperts beschouwden de 'taalkundig naïeve' luisteraars de spraakfragmenten van maar twee matched-guise sprekers als representatief voor de drie accenten. Dit suggereert dat representatieve matched-guises voor onderzoek naar moedertaal en niet-moedertaal Engelse accenten kunnen worden geproduceerd, maar dat spraakevaluaties van 'taalkundig naïeve' luisteraars nodig zijn voor een betrouwbare selectie van matched-guise sprekers.

Om inzicht te krijgen in de oordelen van Nederlanders over een Nederlands-Engels accent in vergelijking met standaard moedertaalaccenten in het Engels is het derde onderzoek uitgevoerd (Hoofdstuk 4). In een matched-guise experiment zijn de reacties van 392 Nederlandse luisteraars op een typisch Nederlands-Engels accent, een standaard Brits-Engels en een standaard Amerikaans-Engels accent gemeten. De

luisteraars moesten de sprekers evalueren (*spraakevaluaties*) op hun sociale status (*status*) en het hebben van een prettige persoonlijkheid (*affectiviteit*). In dit experiment werden de spraakevaluatie dimensies status en affect aangevuld met de spraakevaluatie dimensie *dynamiek* (de dynamiek of mate van energie en enthousiasme van de spreker), op basis van onderzoeksresultaten van andere taalkundig evaluatieonderzoek. Naast spraakevaluaties werden ook begripsvragen (*spraakbegrip*) beantwoord over de inhoud van de spraakfragmenten, vergelijkbaar met de drie types begripsvragen gesteld in het eerste experiment (Hoofdstuk 2). Om te weten te komen of de communicatiecontext waarin spraak wordt gehoord een effect kan hebben op de reacties van luisteraars, werden de drie accenten ook in drie communicatiecontexten beoordeeld die belangrijk zijn in internationale professionele interacties, namelijk onderwijs, toerisme en handel: (1) een hoorcollege, (2) een audiotour en (3) een sollicitatie. De resultaten lieten zien dat, in vergelijking met standaard Brits- en Amerikaans-Engelse accenten, een Nederlands-Engels accent geen negatief effect had op de spraakevaluaties in termen van *affectiviteit* of *dynamiek*, maar wel op de *status* van een spreker. Accent had geen effect op begrip, maar communicatie context had dat wel, want het begrip van de inhoud van de spraakfragmenten (*begrijpelijkheid*) en van het doel van de sprekers (*interpretatie*) was beter voor het hoorcollege dan voor de audiotour en de sollicitatie. Communicatiecontext had ook effect op *affectiviteit*: de sollicitatie resulteerde in significant lagere evaluaties van de vriendelijkheid van de spreker in vergelijking met het hoorcollege en de audiotour. Dit betekent dat in Nederland een Nederlands-Engels accent spraakbegrip niet in de weg hoeft te staan, maar dat accent percepties van sprekers wel negatief kan beïnvloeden, als het gaat om de sociale status van een spreker. Verder bleek communicatiecontext invloed te hebben op het begrip en de evaluaties van sprekers. Deze conclusies zijn gebaseerd op Nederlandse luisteraars. In het laatste onderzoek hebben we ons gericht op de vraag of dit ook geldt voor andere niet-moedertaal luistergroepen.

Het laatste onderzoek is een herhaling van het matched-guise experiment beschreven in Hoofdstuk 4 met dezelfde variabelen *(accent; spraakbegrip; spraakevaluaties; communicatie context)*, maar dan met andere luistergroepen. De niet-moedertaal luistergroepen (N=1699) bestonden uit niet-moedertaalsprekers van het Engels afkomstig uit Duitsland (n=617), Spanje (n=540) en Singapore (n=542). Deze luisteraars vertegenwoordigden niet-moedertaalsprekers van het Engels die belangrijk zijn voor Nederlanders in termen van onderwijs, politiek, toerisme en handel. De resultaten lieten zien dat voor alle luistergroepen een typisch Nederlands-Engels

accent niet minder begrijpelijk was dan een standaard Brits of Amerikaans Engels accent. De Duitse luisteraars maakten in hun spraakevaluaties vrijwel geen onderscheid tussen een spreker met een typisch Nederlands-Engels accent en een spreker met een standaard Brits- of Amerikaans-Engels accent. Voor Spaanse en Singaporezen werden sprekers van een typisch Nederlands-Engels accent zelfs positiever geëvalueerd dan sprekers van een standaard Brits- of Amerikaans-Engels accent.

De algemene conclusies van deze dissertatie zijn dat voor niet-moedertaalsprekers van het Engels het Engels met een Nederlands accent net zo begrijpelijk is als Engels met een standaard Brits of Amerikaans accent. Voor moedertaalsprekers van het Brits-Engels is Engels met een Nederlands-Engelse uitspraak goed te verstaan en te begrijpen, mits het een niet al te sterk Nederlands-Engels accent betreft. De onderzochte luisteraarsgroepen blijken wat betreft spraakevaluaties verschillende indrukken te hebben van een Nederlander die Engels spreekt. Moedertaalsprekers van het Brits-Engels en Nederlanders lijken allebei een taalnorm te hanteren die een niet-moedertaal accent associeert met een lagere sociale status, maar niet perse met een minder prettige persoonlijkheid. Duitsers, Spanjaarden en Singaporezen lijken een andere taalnorm te hanteren: zij staan relatief tolerant ten opzichte van een Nederlands-Engels accent. Spanjaarden en Singaporezen kunnen soms zelfs positievere associaties hebben met een Nederlands-Engels accent dan met een standaard Brits of Amerikaans Engels accent. De resultaten van de Britse luisteraars bevestigen enerzijds het traditionele tweedetaalverwervingperspectief: het leren van een vreemde taal is het beste gelukt wanneer leerders taalvariëteiten kunnen produceren die zoveel mogelijk overeenkomen met die van moedertaalsprekers. Nederlanders lijken dit perspectief ook te delen, maar alleen als het gaat om percepties van sociale status, niet als het gaat om affect, dynamisme of begrip. De resultaten van de Duitsers, Spanjaarden en Singaporezen laten anderzijds ook zien dat een LFE taalgemeenschap lijkt te bestaan met de eerder aangegeven kenmerken van LFE sprekers: een goed begrip van verschillende variëteiten van het standaard Engels en van niet-moedertaal Engels. Het is dan ook de vraag of Engels taalonderwijs gericht zou moeten zijn op het leren van een moedertaal variëteit van het Engels, want accent variëteiten van niet-moedertaalsprekers hoeven begrip en percepties van sprekers niet negatief te beïnvloeden.

ACKNOWLEDGMENTS

I have to start by thanking my most amazing partner, Bram. From making sure I could properly concentrate on my writing by taking the kids out to 'Zwembad West' or an indoor playground to encouraging me to keep on going when I was in no mood to wade through endless piles of data, his support has been fundamental to the successful completion of my research and this book. Thank you so very much babe!

To my sweet babies, my son and my daughter, I love you so much.

Of course I am eternally grateful to my supervisors Marinel Gerritsen, Roeland van Hout, and Brigitte Planken. Marinel has been with me since the very beginning and I can only say that it is primarily because of her expertise and professionalism, but also her immense compassion and generosity that this PhD experience has been very educational and pleasant. Roeland, busy as he is and probably forever will be, still decided to join this research project and as a human linguistics and statistics database was definitely able and to my delight also willing to help out with any question or comment I had. He was always welcoming, patient, clear, to-the-point and truly wonderful to have as a supervisor. Last but definitely not least, Brigitte, who was officially my 'copromotor', which does not seem right, because there was nothing 'co' about her contributions. She encouraged me to develop my independence as a researcher and thoroughly think through my writing and content. Thank you!

I would also like to thank everybody at Radboud in'to Languages and especially Liesbet Korebrits and Sylvia van der Weerden, two great 'bosses' who generously supported my decision to pursue my PhD. In addition, I would like to express my gratitude to the Centre for Language Studies and the Language and Communication Department for giving me 0.2 and 0.3 FTEs of research time for four years to complete the bulk of this research project. Carlos Gussenhoven, thank you for your contributions to the early stages of my PhD. Thank you to my speech sample speakers; you know who you are! My lovely colleagues at the Language and Communication Department, especially Marieke Hoetjes and Frank van Meurs, thank you for a wonderful working environment.

To my family and friends, thank you for your love, support, real talk, joy and fun, I love you all.

CURRICULUM VITAE

Warda Nejjari was born on 18 March 1980 in Nijmegen, the Netherlands. After completing her master degrees in English Language and Literature and International Business Communication, she worked for three years as a business communication lecturer and a communication freelancer for national and regional Dutch governments. From 2009 to 2016 she worked as an English language and communication trainer and programme manager for Radboud University's language and communication centre, Radboud in'to Languages. Currently, she is a lecturer in the Language and Communication Department of Radboud University's Faculty of Arts.