



Nativeness perceptions and speaker voice as predictors of (non-)native English speaker evaluations in four ELF contexts

Warda Nejari*, Roeland van Hout, Marinel Gerritsen, Brigitte Planken

Radboud University, Erasmuslaan 1, 6500HD Nijmegen, the Netherlands

Received 30 June 2023; revised 9 April 2024; accepted in revised form 9 April 2024;

Abstract

We investigated the extent to which responses ($N = 6617$) by four L2 English listener groups (The Netherlands: $n = 1701$; Germany: $n = 1606$; Spain: $n = 1647$; Singapore: $n = 1663$) were affected by giving L1 English speaker status to standard L1 British and American English accents, compared with a typical Dutch English accent. We assessed the extent to which *presumed nativeness* impacted speaker evaluations (status, affect, dynamism), and the extent to which a speaker's *voice* influenced speaker evaluations by analyzing listener responses to verbal and matched guises. The results showed that presumptions of a speaker's nativeness significantly impacted speaker evaluations on all dimensions, and we therefore conclude that speaker evaluations are also based on listeners' views on a speaker's nativeness. In addition, speaker evaluations were influenced by a speaker's voice to such an extent that this can lead to significantly more positive/negative speaker evaluations of both L1 and L2 English speakers. Finally, this study confirms the relevance and main benefit of the matched-guise technique in accentedness research, compared with the verbal-guise technique, since the former successfully minimizes the actual impact of voice.

© 2024 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Keywords: Accentedness; Presumed nativeness; Voice quality; Speaker evaluations; ELF

1. INTRODUCTION

Speaker evaluation research focusing on non-native English accents shows that – in general – (strong) non-native English accents, compared to native English accents, evoke more negative speaker evaluations in terms of social status and (perceived) speech comprehension among native English listeners. However, this does not necessarily apply to the affect felt towards a speaker or listeners' perceptions of a speaker's dynamism (e.g., Bresnahan et al., 2002; Cargile and Giles, 1997; Fuertes et al., 2012; Hendriks et al., 2016; Kalin et al., 1980; Levy-Ari and Keysar, 2010; Lindemann, 2003; Munro and Derwing, 1995a, 1995b; Nejari et al., 2012, 2021; Pihko, 1997). In addition, negative perceptions of non-

* Corresponding author.

E-mail addresses: warda.nejari@ru.nl (W. Nejari), roeland.vanhout@ru.nl (R. van Hout), marinel.gerritsen@ru.nl (M. Gerritsen), brigitte.planken@ru.nl (B. Planken).

native English speaker groups by native English speakers have been proven to invoke negative perceptions of the former's trustworthiness, chances for employment and promotion opportunities in organizations (e.g., [Bond and DePaulo, 2006](#); [Carlson and McHenry, 2006](#); [Da Silva and Leach, 2013](#); [Hosoda et al., 2012](#); [Leach and Da Silva, 2013](#)). Native English, or L1 English, in the context of this article, refers to the English speech produced by those who have learned English as a native speaker in countries where it is the first language of most citizens, such as the United Kingdom, the United States, and Australia. Non-native, or L2, English, refers to both the English spoken as a second language in private and public spheres by people in countries with historical colonial relationships with the United Kingdom (e.g., Singapore, India, or Nigeria) and the English spoken by people who learn it as a foreign language, as in, for example, the Netherlands, Germany, Brazil, Morocco, and Japan (see also [Nejjari et al., 2020](#) for a detailed discussion on this definition).

1.1. L2 listener evaluations of L2 speakers

Even though research oriented towards L2 English listeners evaluating L2 English accents is still relatively scarce, it provides interesting contrasts to the abovementioned studies. For example, [Nejjari et al. \(2020\)](#) studied how three L2 English listener groups (German, Spanish, Singaporean) perceived Dutch-accented English compared to standard British and American English accents. This study showed that the three L2 listener groups were not hindered by being unfamiliar with an L2 English accent where their speech understanding or evaluations of a speaker's status were concerned. In addition, the L2 English accent, compared with the L1, standard English accents, evoked similar or more positive affective evaluations in these listener groups. Indeed, other research also shows that when it comes to speech comprehension, unfamiliar L2 accents can be as comprehensible as L1 accents to both L1 and L2 listeners (see, e.g., [McBride, 2011](#); [Perry et al., 2018](#); [Van Engen and Peelle, 2014](#)). What is important to note is that the abovementioned research focuses on relatively fluent, formal language use, spoken at a reasonable speech rate, with limited distractors, such as background noise (see also [Nejjari et al., 2019, 2020, 2021](#)).

Based on the above, it is reasonable to assume for the present study, which focuses specifically on international professional communication contexts, that L2 English accents used in such settings by fluent L2 English speakers can be understood by other fluent L2 English speakers, even when they are not familiar with the L2 English accent being used. Also, L2 English speakers do not appear to apply a language norm that connects L2 English accents to lower status, affect and dynamism, when compared to L1 English accents (see also [Canagarajah, 2007](#)). However, this assumption is based on findings from research on speaker evaluations that focus on accents. Such studies have infrequently investigated the underlying mechanisms of speaker evaluations. The present study has aimed to examine such mechanisms in greater depth.

In a follow-up study by [Nejjari et al. \(2021\)](#), a Dutch listener group that was familiar with Dutch-accented English reacted to the same three accents as in the earlier study ([Nejjari et al., 2020](#)). In this study, the Dutch listeners, unlike their German, Spanish, and Singaporean counterparts, evaluated the status of Dutch-accented English speech more negatively, when compared to L1, standard English. This suggests that when L2 listeners are familiar with a particular L2 English accent, they readily gave that accent L2 status, which evokes specific speaker associations in listeners. These speaker associations in turn lead the listeners to apply a language norm that holds that specific L2-accented English speech is indicative of a speaker with lower status, when compared to L1-accented English speech.

The findings from [Nejjari et al. \(2020, 2021\)](#), taken together, suggest that speaker evaluation processes are at least partly based on categorizing, and thus presuming, the accent origins of speakers, which is supported by other studies (see e.g., [Buckingham, 2014](#); [Dragojevic et al., 2018](#); [McKenzie, 2008](#); [Yook and Lindemann, 2013](#)). For instance, [McKenzie \(2008\)](#) investigated L1 Japanese speakers' speaker evaluations of four L1 English accents, as well as one strong and one weak Japanese-English accent: he demonstrated that correctly categorizing, and thus presuming, L1 English speech as being L1 English led to more positive status evaluations of L1 English speakers, but that correctly categorizing accents was not necessarily easy for listeners. Furthermore, [Yook and Lindemann \(2013\)](#), studied Korean L2 English listeners' ability to identify L1 and L2 English accents and how their awareness of these accents' origins influenced their related evaluations. Generally, the listeners preferred standard, L1-accented English. However, when listeners then had to respond to L1 and L2 English accents, only those listeners who had been made aware of the tested accents' origins displayed speaker evaluations that followed their previously indicated preference for an L1, standard English accent. The listeners who had not been made aware of the speakers' accent origins did not display these speaker evaluation patterns, indirectly suggesting that a listener's ability to categorize specific accents correctly cannot be assumed. This is also supported by other research that shows that, beyond general distinctions between L1 and L2 accents, or a standard or non-standard accent, listeners often struggle to categorize the exact accent they are listening to correctly, unless they are familiar with such an accent ([Brunner, 2009](#); [Gnevshva, 2018](#); [McGorrrery and McMahon, 2017](#); [Nejjari et al., 2012, 2020, 2021](#); [Nolan, 2003](#); [Wong and Babel, 2017](#)).

Moreover, recent studies also suggest that the human brain, in interaction with our senses, attempts to process and foresee the makeup and development of its surroundings (see e.g., [De Lange et al., 2018](#); [Miyamoto et al., 2021](#)), as well as when dealing with language and its speakers (see, e.g., [Kajiura et al., 2021](#)). By extension, this perspective suggests that when we listen to someone, we apply our language knowledge and proficiency to predict what that speaker will say. Furthermore, as evolutionary psychology research implies, we use accent as a cue to predict a speaker's language community, for example, one with an L1 or L2 status (see, e.g., [Pietraszewski and Schwartz, 2014](#)).

In the context of English as a global lingua franca, one can imagine that the precise accent origins of L2 English speakers are often not known when (professional) interactions between L2 English speakers take place. It is therefore relevant to investigate how L1 and L2 English accent origins are perceived by different L2 English listener groups, as this might allow us to better understand the extent to which giving someone (non-)nativeness status impacts speaker evaluations in professional settings.

From a practical perspective, if L2 English listeners giving L1 or L2 status to speakers of English does not always impact speaker evaluations, this might mean that English language (teaching) training does not have to focus on making learners and teachers come across as native-like in English as possible. Most English speakers today are in the L2 category and English is most frequently used in international professional contexts. Consequently, English language training should perhaps be focused on effective communication between L2 English speakers, on the basis of a functionally sufficient (but not necessarily native or native-like) level of adequate English language competence, suitable for the specific professional context (see also [Jenkins, 2009](#); [Zoghbor, 2018](#)).

1.2. Voice

Furthermore, to help establish the precise role that giving (non-)native status to speakers plays in speaker evaluation processes, it is also important to understand whether other speech characteristics, such as a speaker's voice, influence speaker evaluations. A speaker's voice communicates important indexical information, such as their gender, age and emotional state, with voice having the ability to express approximately 24 emotions (e.g., [Majid, 2012](#)), regardless of the language used (e.g., [Cowen et al., 2019](#); [Kreiman et al., 1993](#); [Pell et al., 2009](#)). Moreover, an individual's voice can vary on the basis of cultural background and communication context, and voice use also changes on the basis of relationships people have with specific individuals and groups. For example, intimate relationships lead people to use their voice differently compared with formal, professional relationships (e.g., [Ding et al., 2018](#); [Pajupuu et al., 2019](#); [Yiu et al., 2008](#)).

The impact of a speaker's voice on speaker evaluations can be investigated by conducting experiments that include both verbal and matched guises. Verbal guises are audio stimuli produced by L1 speakers of the selected accents, and matched guises are audio stimuli in various selected accents produced by one speaker. The use of matched guises (as distinct from verbal guises) is believed to better ensure that responses to accents are not based on an individual speaker's voice (see further discussion in [Nejjari et al., 2019](#)), but on the produced accents that, like verbal guises, need to sound authentic as well as natural (see [Garrett, 2010](#); [Nejjari et al., 2019](#)). Matched guises work very well in accentedness research because listeners tend not to be very skilled at recognizing the same voice in different accents (see [McGorrey and McMahon, 2017](#) for an overview).

The reality of the use of English as a lingua franca raises the likelihood that people do not know each other's linguistic backgrounds, and often involves communication with people who have unfamiliar accents. It is therefore relevant to explore the process of speaker evaluations and the mechanisms that potentially influence it, such as presumed (non-)nativeness and voice. To explore the latter, we used data collected by [Nejjari et al., \(2020,2021\)](#) in the Netherlands, Germany, Spain and Singapore to conduct a cross-national comparative study that consisted of reactions to both verbal and matched guises. [Nejjari et al., \(2020,2021\)](#) collected data on reactions to matched and verbal guises of L1, standard British and American English accents and Dutch-accented English, but used the matched guises only to compare responses to the selected accents, to minimize the impact of voice. L1 English in the context of this study refers to the accents produced by British and American speakers of English (verbal and matched guises). The L2 English listeners in this study were Dutch, Spanish, German and Singaporean (See also 1. Introduction).

Both the verbal guises and the matched guise were used as stimuli to investigate the extent to which L2 English listener groups give L1 and L2 English accent varieties L1 and L2 status (*presumed nativeness*), and how this affects speaker evaluations (status, affect, dynamism). Furthermore, we included analyses of reactions to individual speakers' voices to better understand the extent to which speaker evaluation processes are also impacted by a speaker's voice. This made it possible to address the following research questions:

- 1) To what extent do L2 English listener groups associate standard British and American English accents and Dutch-accented English with (non-)nativeness, and to what extent do these associations impact speaker evaluations?

2) To what extent does a speaker's voice impact L2 English listeners' speaker evaluations?

2. METHOD

We measured the responses ($N = 6617$) from four listener groups from the Netherlands ($n = 1701$), Germany ($n = 1606$), Spain ($n = 1647$) and Singapore ($n = 1663$) to both matched and verbal guises. All the listeners responded to stimuli (four speech samples per listener) via an online questionnaire. The experiment had a within-subject, multi-factorial design. All the listeners (*listener groups*) were exposed to the independent variables (*accent, voice*) and evaluated the stimuli on the dependent variables (*presumed nativeness, speaker evaluations*).

2.1. Speakers

The responses to eight speakers of English (*voice*) were included in the current study. One male matched-guise speaker (MG) produced the three accents that represented *accent*: (1) standard British English, (2) standard American English, and (3) the typical English accent of highly educated L1 speakers of Dutch (see Nejari et al., 2019, 2020, 2021). The standard accents of British and American English refer in this study to those generally associated with the national accent norm of these nations and are generally similar to what is commonly referred to as a standard British English accent and General American for a standard American English accent. A typical Dutch English accent in the present study is defined as having features that L1 speakers of Dutch and others familiar with Dutch and Dutch English will recognize as such (see Nejari et al., 2019, 2020 for further description of phonological features of Dutch-accented English).

We also included verbal-guise speech samples by six male speakers (V1-6): two L1 Dutch speakers with a typical Dutch accent in English, two L1 speakers of standard British English, and two L1 speakers of standard American English. One filler speech sample in a standard British English accent was presented to all listeners at the beginning of the experiment to familiarize listeners with the task, create a benchmark with which listeners could compare the other speech samples, and establish whether the listener responses were consistent. All the speakers were aged 35–60 at the time of recording; had at least a master's degree; were English language and/or linguistics specialists, and teachers/ lecturers; and were therefore used to speaking in public in a professional capacity (see Table 1). Voice research proposes that healthy adults' voices change as they age, for example, in terms of fundamental frequency, but also that listener perceptions can significantly impact perceptions of voice characteristics (Eichhorn et al., 2018; Lortie et al., 2018; Rojas et al., 2020). However, due to a lack of data, variability in research methods applied, and limited methodological information, it remains challenging to determine conclusively the extent to which the human voice changes, and precisely when and why this happens (see meta-analysis in Rojas et al., 2020). Voice research does appear to suggest that healthy adults between approximately 30–60 years of age are unlikely to differ significantly in their acoustic voice properties. Therefore, the speech production of speakers in this age range, as with the speakers in this study, can be regarded as representing the middle-aged adult human voice (see Eichhorn et al., 2018). Also, on the basis of the fact that our speakers were professional speakers, it was assumed that the samples they produced, even in an experimental setting, would be as natural and authentic as possible.

All but verbal-guise speaker six (V2) had been selected on the basis of a speech sample experiment conducted by Nejari et al. (2019). This experiment employed intraclass correlation coefficients on reactions by L1 listener groups in the Netherlands, the United States and the United Kingdom to determine speech samples that were representative for the three accents. In Nejari et al. (2019), both the matched-guise and the verbal-guise speakers' speech samples in the selected accents were evaluated on their degree of 'nativeness' and 'standardness' by three 'linguistically naive' listener groups: 40 native British English speakers who evaluated the British English samples, 40 native American English speakers who evaluated the American English samples, and 40 L1 Dutch speakers who evaluated the Dutch English samples. The Dutch-accented English speech samples were evaluated on nativeness and typicalness (instead of stan-

Table 1
Speakers per accent.

Speech samples per accent	One matched-guise speaker	Six verbal-guise speakers	One filler speaker
Dutch English	MG	V1, V2	
Standard British English	MG	V3, V4	Filler
Standard American English	MG	V5, V6	

dardness) to assess whether the selected speakers produced speech samples that were representative of the selected Dutch-accented English.

The matched-guise speaker had been part of a group of four matched-guise speakers that were selected by the researchers to produce the three accents. These speakers were all linguists and English language specialists with specific linguistic expertise in terms of Dutch/British/American English (accents). Nejari et al. (2019) selected these four speakers to better guarantee that reactions to the speech samples were based on their accent and not personal characteristics such as voice quality and speech rate: it was considered that such characteristics can result in a more positive/negative evaluation of an individual speaker on the basis of perceptions of those 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. The matched-guise speaker who was considered as being a representative speaker for the three selected accents was selected for our follow-up accentedness experiments, including the current study. Verbal-guise speaker number two had been informally assessed by three experienced linguistics professors to represent a typical Dutch-English accent.

2.2. Instrumentation

Stimuli developed by Nejari et al., (2020,2021) were used in the current study. One filler text (on technological development and robotics, taken from Nejari et al., 2019), and three stimulus texts (marketing lecture, art gallery audio tour, job pitch) that represented international professional communication contexts in which English is commonly used as a lingua franca (see stimuli in Supplementary Materials 1 via: <https://surfdrive.surf.nl/files/index.php/s/h8M2-tan2AaZmRSt>). Although communication context was a factor in the earlier research by Nejari et al., (2020,2021), it was not of interest in the present study.

The matched-guise and verbal-guise speakers produced the three accents in all three contexts, resulting in nine speech samples by the matched-guise speaker (3 accents x 3 contexts), and 18 samples by the verbal-guise speakers (6 speakers x 3 contexts). The filler speaker produced one speech sample in standard British English. The speech samples were between 40 and 60 s long. Each listener evaluated four different speech samples by four different speakers: the filler sample followed by three different accent samples (in this order: Lecture, Audio Tour, Job Pitch) produced by the matched-guise speaker and the verbal-guise L1 speakers. To ensure that the speech samples could be evaluated for each accent and speaker, to avoid repeating the content of each context, and to limit any order effect, each listener group was split into 18 subgroups, with a targeted 30 listeners per listener group in each country (Table 2). This resulted in at least 540 listeners per listener group (Table 3).

On the first page of the questionnaire, the listeners were provided with a general introduction and were told that this was a study focused on international professional communication, but were not informed of the purpose of the study. To ensure that listeners would provide their first impression of the speakers' traits (*speaker evaluations*), all listeners answered the *speaker evaluation* questions first, by clicking on a link to the speech sample being evaluated, listening to it, and then answering the question. The listeners were then asked to indicate where they believed a speaker was from, that is, the speaker's country of origin (see, e.g., Dragojevic et al., 2018; Hendriks et al., 2016; McKenzie, 2008). This is a common, general question that researchers use for assessing the extent to which speech is associated with a country and the main – or even official – language that is associated with that national identity: for example, French for speakers perceived to come from France, Mandarin Chinese for China, and English for the UK. We also used this question to avoid influencing listeners' behavior by asking questions that could reveal the actual intent of the experiment. We were primarily interested in the associations people have with the accent varieties they listened to, to assess the extent to which this impacted speaker evaluations, not their actual ability to identify them correctly.

The answers to the country of origin question were used to determine what we have defined as *presumed nativeness*, which means that we tested the impact of a listener presuming that they are listening to an L1/2 speaker of English and the effects of this presumption on speaker evaluations. A speaker perceived to originate from the UK, the US, Canada, Australia, New Zealand, South Africa or Ireland was categorized as having L1 English speaker status, while others were not. As described in section 3.1, all listeners overwhelmingly categorized the standard British and American English accents by the matched and verbal guises as being of UK and US origin, suggesting that they associated the tested L1 English accents with L1 English speaking countries and therefore with L1 English. The listeners also answered comprehension questions for all speech samples.¹ All the comprehension or *speech understandability* scores

¹ Three questions were used to measure comprehension: (1) listeners had to listen to a segment from the speech sample and were asked to write down what was literally stated; (2) listeners were asked to indicate whether statements on the topic of the samples were correct or not; (3) listeners were asked to indicate whether statements on the communicative intentions of the speakers were correct or not.

Table 2
Subgroups per listener group (the Netherlands, Germany, Singapore, Spain).

N*	Version 1		Version 2
1 <i>n</i> = 30	**Filler MG BrE Lecture V5 AmE Audio Tour V3 BrE Job Pitch	10 <i>n</i> = 30	Filler MG AmE Lecture V6 AmE Audio Tour V4 BrE Job Pitch
2 <i>n</i> = 30	Filler V1 DE Lecture MG BrE Audio Tour V5 AmE Job Pitch	11 <i>n</i> = 30	Filler V2 DE Lecture MG AmE Audio Tour V6 AmE Job Pitch
3 <i>n</i> = 30	Filler V5 AmE Lecture V1 DE Audio Tour MG BrE Job Pitch	12 <i>n</i> = 30	Filler V6 AmE Lecture V2 DE Audio Tour MG AmE Job Pitch
4 <i>n</i> = 30	Filler MG AmE Lecture V6 AmE Audio Tour V1 DE Job Pitch	13 <i>n</i> = 30	Filler MG DE Lecture V5 AmE Audio Tour
5 <i>n</i> = 30	Filler V3 BrE Lecture MG AmE Audio Tour V6 AmE Job Pitch	14 <i>n</i> = 30	V2 DE Job Pitch Filler V4 BrE Lecture MG DE Audio Tour V5 AmE Job Pitch
6 <i>n</i> = 30	Filler V6 AmE Lecture V3 BrE Audio Tour MG AmE Job Pitch	15 <i>n</i> = 30	Filler V5 AmE Lecture V4 BrE Audio Tour MG DE Job Pitch
7 <i>n</i> = 30	Filler MG DE Lecture V2 DE Audio Tour V4 BrE Job Pitch	16 <i>n</i> = 30	Filler MG BrE Lecture V1 DE Audio Tour V3 BrE Job Pitch
8 <i>n</i> = 30	Filler V4 BrE Lecture MG DE Audio Tour V2 DE Job Pitch	17 <i>n</i> = 30	Filler V3 BrE Lecture MG BrE Audio Tour V1 DE Job Pitch
9 <i>n</i> = 30	Filler V2 DE Lecture V4 BrE Audio Tour MG DE Job Pitch	18 <i>n</i> = 30	Filler V1 DE Lecture V3 BrE Audio Tour MG BrE Job Pitch

N* = 540; **Filler = filler speaker; MG = matched-guise speaker; V 1,2,3,4,5,6 = verbal-guise speaker 1,2,3,4,5,6; BrE = standard British English; *AmE = standard American English; ****DE = Dutch-accented English.

were high, with little variation between listener groups, and hence these did not impact *speaker evaluations* (see Nejari et al., 2021, 2020 for *speech understandability* results). Finally, listeners were asked to answer questions regarding their English language skills, what L1 languages they spoke, and biographical details.

Table 3

Listeners ($N = 2266$): age, % sex, self-reported English fluency, education level.

	Netherlands ($N = 567$)	Germany ($N = 617$)	Spain ($N = 540$)	Singapore ($N = 542$)
Mean age:	38 (Min = 18; Max = 83)	38 (Min = 19; Max = 83)	37 (Min = 18; Max = 64)	34 (Min = 18; Max = 80)
Male:	39.7%	47.9%	36.1%	44.6%
Female:	60.3%	52.1%	63.9%	55.4%
Mean self-reported English fluency (SD) (Min = 1; Max = 5) ^a :	3.61 (SD = .59) (Min = 2.00; Max = 5.00)	4.03 (SD = .51) (Min = 2.75; Max = 5.00)	3.62 (SD = .50) (Min = 2.25; Max = 5.00)	4.21 (SD = .69) (Min = 2.00; Max = 5.00)
Education: A-level	12.2%	0.0%	0.0%	0.2%
Undergraduate/Bachelor:	59.5%	42.1%	55.0%	84.4%
Master:	19.5%	52.0%	40.5%	12.3%
Doctorate:	2.4%	5.8%	4.5%	2.0%
Other:	6.4%	0.0%	0.0%	1.1%

^aMean self-reported English fluency was the mean for indicated levels for English listening, reading, writing, speaking skills on a 5-point scale. (1: very low; 2: low; 3: average; 4: high; 5: like an L1 speaker).

2.3. Data collection

The data were collected online in 2016, 2017, and 2018, mostly via Qualtrics, a global survey software and online data collection company that caters for (non-) commercial organizations. Qualtrics was hired to sample listeners who were L1 speakers of the main national language of the Netherlands (Dutch), Germany (German), Spain (Spanish), and Singapore (including the national language, Malay, and the officially-acknowledged English, with officially-acknowledged languages like Mandarin and Tamil also accepted): the sample listeners also needed to be highly educated (i.e., having at least attained or completed undergraduate level education). In the Netherlands, approximately 30% of data were also collected via a Facebook page, but the majority via Qualtrics. Of the German listener data, only 5% was initially collected in the context of a research course at Radboud University in the Netherlands, with the remaining 95% collected via Qualtrics. All the listener data from Spain and Singapore was collected via Qualtrics.

Data collection took place in three rounds for all countries. With each round, incomplete questionnaires (approximately 25%), and nonsense answers or symbols and/or only neutral answers (approximately 15%) that listeners provided were replaced by new responses from other listeners. The median time to complete the questionnaire was just over 16 min for all listeners. No restrictions were placed in terms of regions in the four countries where the experiment was conducted. Table 3 shows the number of completed questionnaires per listener group, the listener groups' mean age, sex, self-reported English fluency, and education level.

2.4. Speaker evaluations and presumed nativeness

Speaker evaluations and the listeners' estimation of the speakers' *presumed nativeness* (Q: 'Where do you believe the speaker is from?') were assessed for each speaker. Based on Nejari et al., (2012,2020,2021), speaker evaluations were measured by asking listeners to indicate on 5-point Likert scales (1 = strongly disagree; 5 = strongly agree; 3 = neither disagree nor agree) the extent to which they believed the speaker possessed 11 personality traits, representing three dimensions: *status* (competent, educated, having authority, intelligent, cultured), *affect* (considerate, pleasant, friendly), and *dynamism* (energetic, enthusiastic, confident). The three dimensions are commonly employed in speaker evaluation research to gauge listener perceptions of a speaker's intelligence and competence (*status*), a speaker's likeability (*affect*), and their self-presentation (*dynamism*) (see, e.g., Grondelaers and Van Hout, 2015; Zahn and Hopper, 1985). To confirm the dimensionality of the speaker evaluation items, a principal component analysis was applied with an Eigenvalue > 1 criterion for factor extraction, and varimax rotation. The personality items showed a resolution into three factors: *status*, *affect*, *dynamism*, as can be seen in Table 4.

Table 4

Rotated Factor Matrix: factor loadings scores on 11 scales with three factors for listener groups (Netherlands, Germany, Spain, Singapore). Only loadings > .600 are reported.

	Factor 1	Factor 2	Factor 3
Educated	.825		
Intelligent	.780		
Cultured	.780		
Competent	.723		
Confident	.610		
Friendly		.843	
Pleasant		.725	
Considerate		.709	
Energetic			.825
Enthusiastic			.657
Authoritative			.680

Status was analyzed for the personality traits *educated*, *intelligent*, *cultured*, *competent*, and *confident*. *Affect* was analyzed for the personality traits *friendly*, *pleasant*, and *considerate*, because all listeners evaluated these personality traits as part of one factor. *Dynamism* was analyzed for the personality traits *energetic*, *enthusiastic* and *authoritative*. As a result, three factors were defined and used in further analyses: *status* (educated, intelligent, cultured, competent, confident), *affect* (friendly, pleasant, considerate), and *dynamism* (energetic, enthusiastic, authoritative).

2.5. Statistics

Descriptives and frequencies were calculated to establish means and percentages of listener characteristics and responses (Table 3). The factor analyses we applied to trace underlying dimensions in speaker evaluations were principal component analyses, with varimax rotation (SPSS, release 27). We used R package lme4, to carry out the mixed regression analyses, with participant as a random factor. We scrutinized residual scores to see whether we had serious violations of underlying statistical assumptions. There were no compelling reasons, given the robustness of both techniques, to apply alternative statistical techniques.

3. RESULTS

Firstly, we shall describe the extent to which the four listener groups (Dutch, German, Spanish, Singaporean) gave the matched- and verbal-guise speakers' L1 English speaker status (*presumed nativeness*) (3.1). Secondly, we shall discuss the impact of perceiving speakers as L1 and L2 English speakers (*presumed nativeness*) on *speaker evaluations* (*status*, *affect*, *dynamism*) (3.2). Lastly, we shall discuss listeners' responses to the individual speakers (*voice*) to understand the effects of both *voice* and *accent* on *speaker evaluations* and *presumed nativeness* (3.3).

3.1. Presumed nativeness: Who is the real L1?

Fig. 1 shows the percentages of times the matched-guise (MG) and verbal-guise (V1,2,3,4,5,6) speakers were presumed to be L1 English speakers (*presumed nativeness*), per listener group (country). Table 5 in Supplementary Materials 2 via <https://surfdribe.surf.nl/files/index.php/s/zRXvo2KG3HRaRgJ> shows the specific country of origin answers that were given per speaker and listener group in percentages. As explained in 2.1, all but one of the speakers (V2 was approved by three linguistics professors) had been assessed in a previous experiment to ensure representative speech samples for all three accents (Nejjari et al., 2019). This means that to L1 speakers of each accent variety, all the selected matched and verbal guises were easily recognizable as representing the three tested English accents.

Fig. 1, with substantial overlap in confidence intervals between most scores for the L1 English guises, shows that the matched-guise speaker was able to produce L1 English accents convincingly, according to all listener groups (matched guises: standard British English > 84% within each listener group; standard American English > 86% within each listener group). An overlap in confidence intervals means that there are no significant differences. In addition, the majority of these listeners also correctly indicated that the standard British and American English matched guises were British and American English accents. The verbal-guise speakers of the L1 English accents were also deemed convincing as

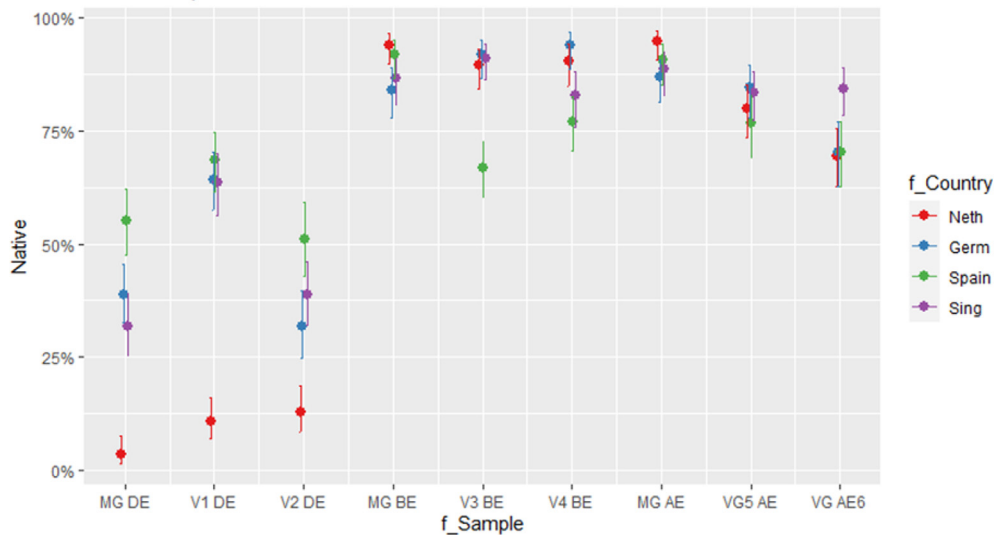


Fig. 1. Presumed nativeness in %, per speaker, listener group*, 95% confidence intervals. ***f_Sample**: MG = matched-guise speaker; V1,2,3,4,5,6 = verbal-guise speaker 1,2,3,4,5, or 6; DE = Dutch-accented English; BE = standard British English; AE = standard American English; **f_Country**: Neth = the Netherlands, Germ = Germany, Spain = Spain, Sing = Singapore.

L1 English speakers, according to all listener groups, with the standard British English verbal guises somewhat more so than the standard American English verbal guises (standard British English > 82%; standard American English > 69%).

As shown in Fig. 1, The Dutch-accented English matched guises were presumed to be L1 English speakers to some degree by all listener groups – to the greatest degree, given the non-overlapping confidence intervals, by Spanish listeners and to the least by Dutch listeners (L2 > 94% of listeners). Fig. 1 also shows substantial confidence interval overlap in the *presumed nativeness* scores for the Dutch-accented matched and verbal guises, meaning that the Dutch listeners viewed all guises similarly. Most German and Singaporean listeners also believed they were listening to L2 English speech; but at least 30% of them, and 55% of Spanish listeners, believed they were listening to L1 English speech. In those cases where listeners believed they were listening to an L1 English speaker, they mainly indicated that the speaker was from either the United Kingdom, the United States or Australia, and not from another L1 English speaking country. Where listeners believed they were listening to L2 English, the countries of origin most often indicated were the Netherlands for Dutch and German listeners, and France for Spanish and Singaporean listeners.

Similar to the Dutch-accented matched guises (94.2% for MG DE), the Dutch-accented English verbal guises were recognized best by Dutch listeners, albeit to a lesser degree when compared with the matched guises, with over 10% of Dutch listeners believing they were listening to an L1 English speaker. German, Spanish, and Singaporean listeners clearly distinguished between the verbal-guise speakers (V1,2 DE) in terms of nativeness. For example, over 63% of German, Spanish, and Singaporean listeners viewed verbal-guise speaker 1 as an L1 English speaker. Verbal-guise speaker 2 was viewed by at least 30% of German and Singaporean listeners as L1 English and 55% of Spanish listeners as L1 English: a pattern similar to that found for our Dutch-English matched-guise speaker. In general, Spanish listeners had the most difficulty distinguishing between L1 and L2 English accents. The verbal guises viewed as L1 English were viewed as originating either from the United Kingdom, the United States, or Australia, and not other L1 English speaking countries. The verbal guises not considered to be L1 English were mostly assigned Dutch or European origins. Interestingly, some Singaporeans (5.6%) assigned Indian origins to the Dutch-accented matched guises. Similarly, in 10% – 16% of cases, German, Spanish, and Singaporean listeners believed they were listening to Indian-accented English when listening to verbal-guise speaker two (V2).

3.2. The impact of presumed nativeness on speaker evaluations

Mixed model analyses were conducted to see the extent to which the *speaker evaluations* (*status*, *affect*, *dynamism*) were impacted by *presumed nativeness*, and whether this differed per *listener group*. These results are reported below. The means of the *speaker evaluations* and the frequencies for *presumed nativeness* per speaker, accent, and listener

group are given in Figs. 2–7 (see Table 6 in Supplementary Materials 3 via <https://surfdribe.surf.nl/files/index.php/s/fFMZCCrttF7xIGC> for means and SDs).

To visualize the mixed model analyses, Figs. 2–7 were included. They show how the tested speaker evaluation dimensions (status, affect, or dynamism) were impacted by listener groups (country). The figures were split into figures of the responses by listeners who had given speakers the status of L2 English speaker (Figs. 2, 4, 6), or that of an L1

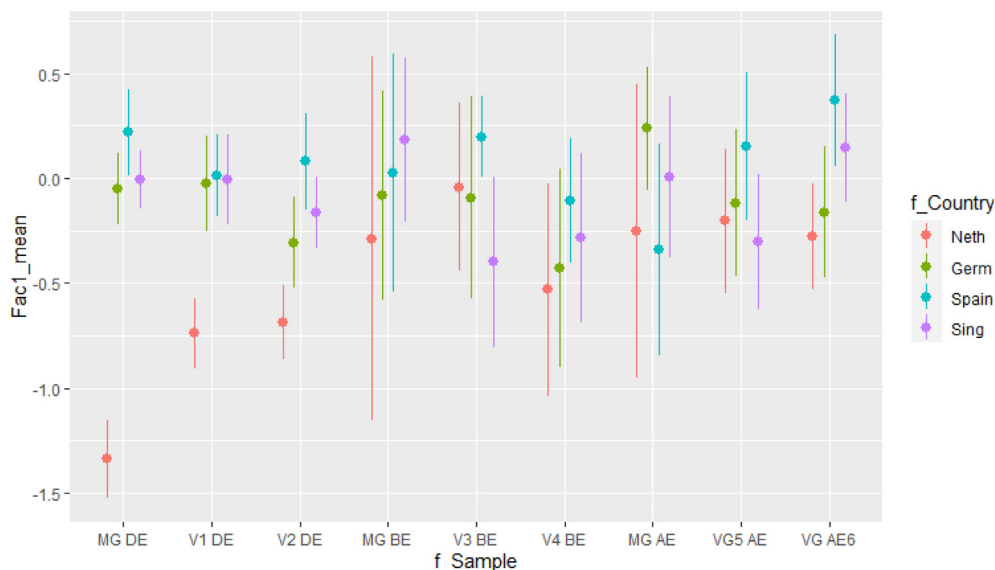


Fig. 2. Status factorial means (Fac1_mean) for speakers' presumed non-nativeness, per listener group (f_Country)*, 95% confidence intervals. *Fac1_mean = Status; f_Sample: MG = matched-guise speaker; V1,2,3,4,5,6 = verbal-guise speaker 1,2,3,4,5, or 6; DE = Dutch-accented English; BE = standard British English; AE = standard American English; f_Country: Neth = the Netherlands, Germ = Germany, Spain = Spain, Sing = Singapore.

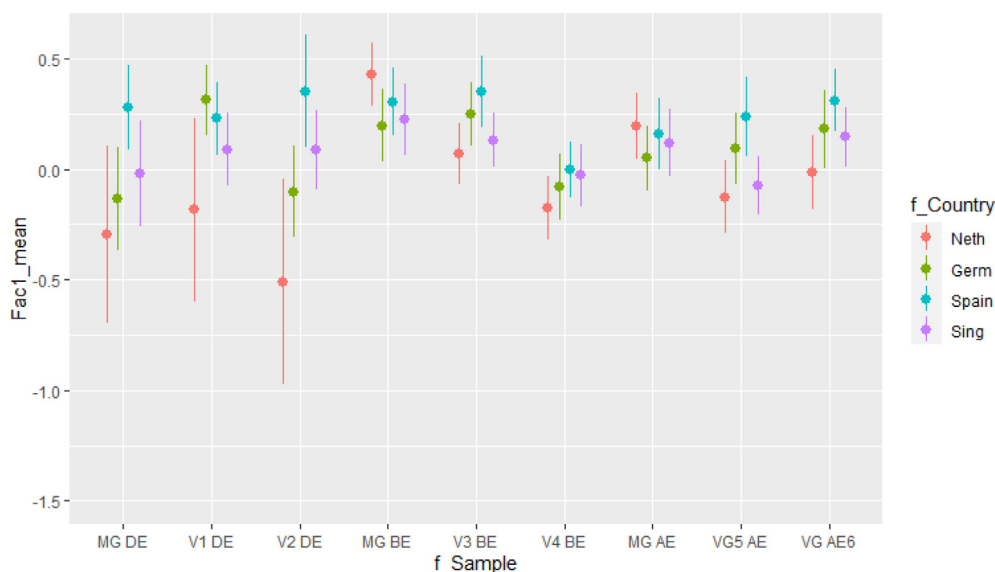


Fig. 3. Status factorial means (Fac1_mean) for speakers' presumed nativeness, per listener group (f_Country)*, 95% confidence intervals. *Fac1_mean = Status; f_Sample: MG = matched-guise speaker; V1,2,3,4,5,6 = verbal-guise speaker 1,2,3,4,5, or 6; DE = Dutch-accented English; BE = standard British English; AE = standard American English; f_Country: Neth = the Netherlands, Germ = Germany, Spain = Spain, Sing = Singapore.

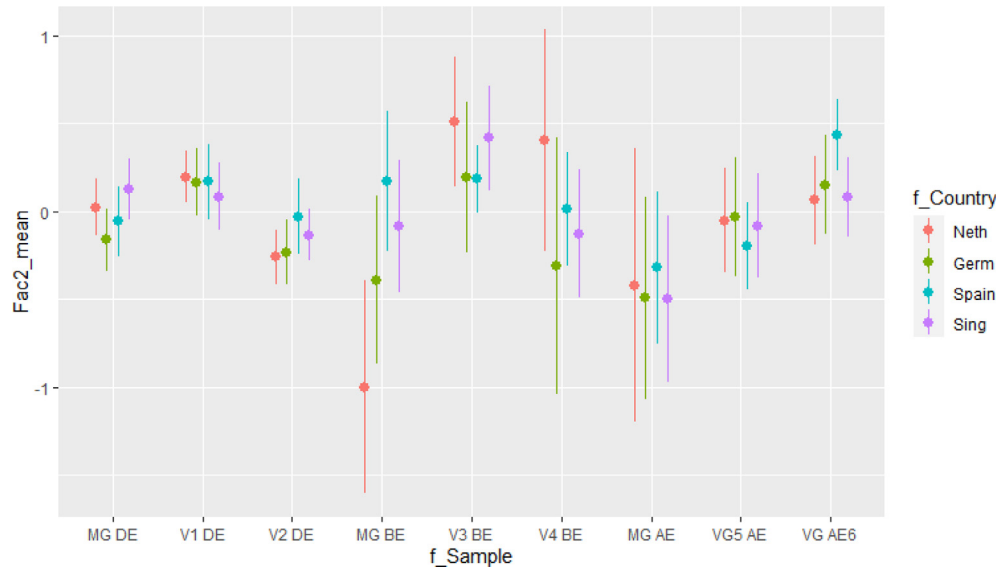


Fig. 4. Affect factorial means (Fac1_mean) for speakers' presumed non-nativeness, per listener group (f_Country)*, 95% confidence intervals. ***Fac2_mean** = Affect; **f_Sample**: MG = matched-guise speaker; V1,2,3,4,5,6 = Verbal-guise speaker 1,2,3,4,5, or 6; DE = Dutch-accented English; BE = standard British English; AE = standard American English; **f_Country**: Neth = the Netherlands, Germ = Germany, Spain = Spain, Sing = Singapore.

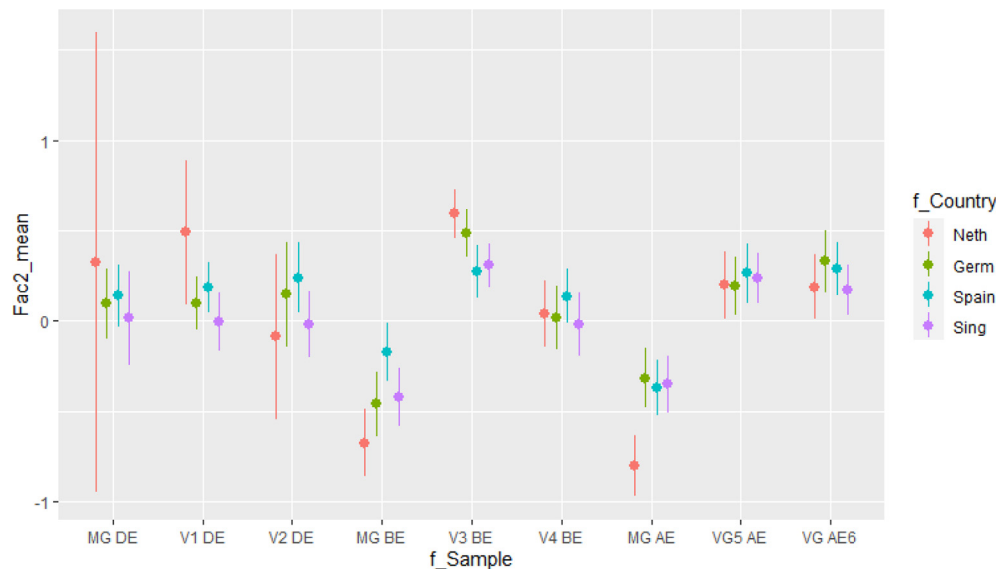


Fig. 5. Affect factorial means (Fac1_mean) for speakers' presumed nativeness, per listener group (f_Country)*, 95% confidence intervals. ***Fac2_mean** = Affect; **f_Sample**: MG = matched-guise speaker; V1,2,3,4,5,6 = verbal-guise speaker 1,2,3,4,5, or 6; DE = Dutch-accented English; BE = standard British English; AE = standard American English; **f_Country**: Neth = the Netherlands, Germ = Germany, Spain = Spain, Sing = Singapore.

English speaker (Figs. 3, 5, 7). By displaying the confidence intervals and their means of the observed speaker evaluation, the reported mixed model effects are demonstrated. Non-overlapping confidence intervals show a significant effect.

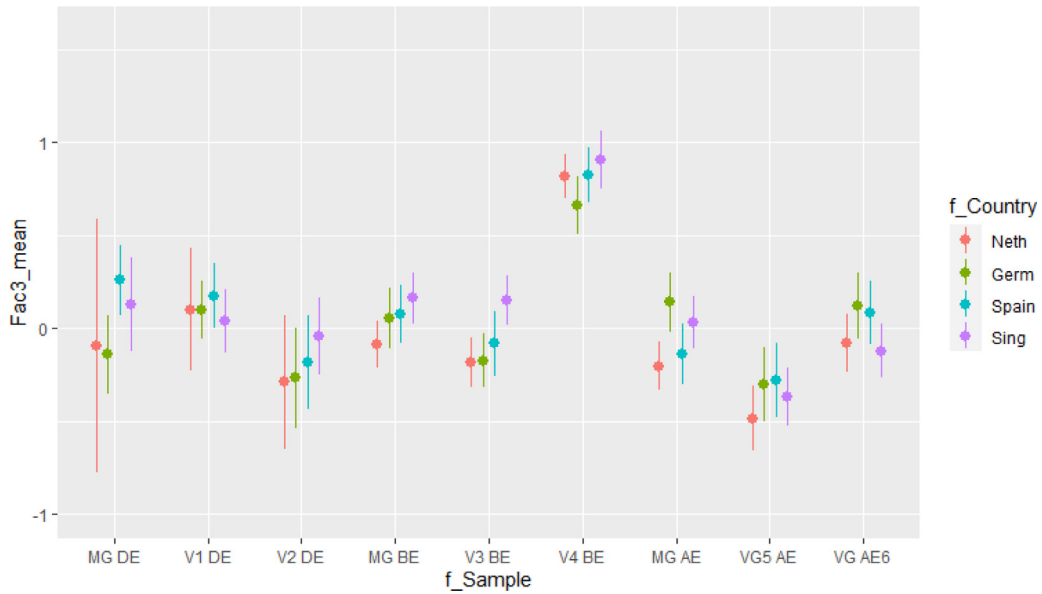


Fig. 6. Dynamism factorial means (Fac1_mean) for speakers' presumed nativeness, per listener group f_Country)*, 95% confidence intervals. ***Fac3_mean** = Dynamism; **f_Sample**: MG = matched-guise speaker; V1,2,3,4,5,6 = Verbal-guise speaker 1,2,3,4,5, or 6; DE = Dutch-accented English; BE = standard British English; AE = standard American English; **f_Country**: Neth = the Netherlands, Germ = Germany, Spain = Spain, Sing = Singapore.

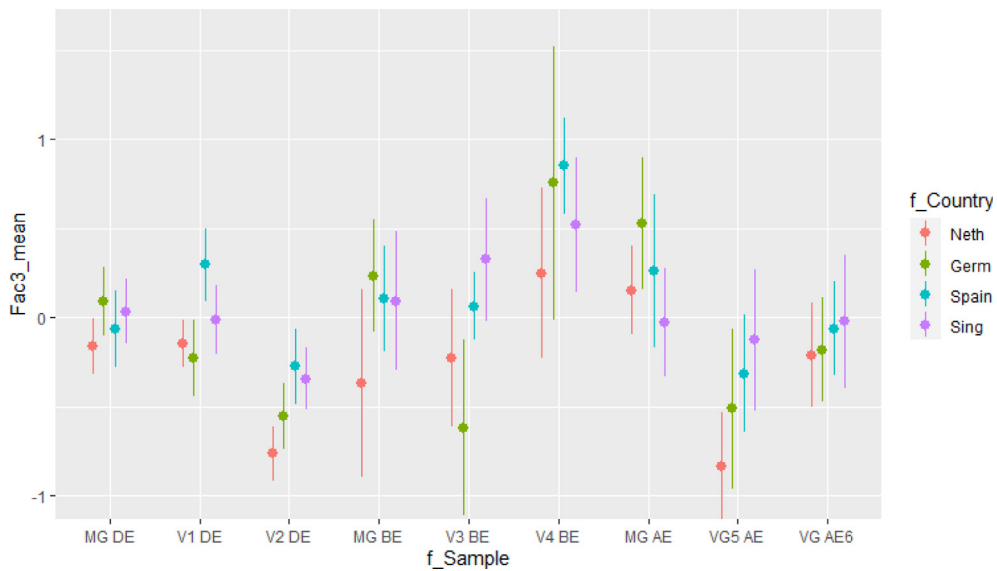


Fig. 7. Dynamism factorial means (Fac1_mean) for speakers' presumed non-nativeness, per listener group f_Country)*, 95% confidence intervals. ***Fac3_mean** = Dynamism; **f_Sample**: MG = matched-guise speaker; V1,2,3,4,5, or 6 = Verbal-guise speaker 1,2,3,4,5, 6; DE = Dutch-accented English; BE = standard British English; AE = standard American English; **f_Country**: Neth = the Netherlands, Germ = Germany, Spain = Spain, Sing = Singapore.

3.2.1. Status and presumed nativeness

A linear mixed model analysis showed that all effects were significant (Figs. 2, 3). There was a three-way interaction between *speaker*, *presumed nativeness*, and *listener group* ($F(18, 6220.4) = 6.78, p = .00$). All three two-way interactions were significant as well: *speaker* and *presumed nativeness* ($F(6, 6193.9) = 3.58, p = .01$), *presumed nativeness*

and *listener group* ($F(3, 6324.1) = 4.32, p = .00$), and *speaker* and *listener group* ($F(18, 6104.7) = 2.82, p = .00$). Finally, there were three main effects: *speaker* ($F(6, 5906.2) = 4.57, p = .00$), *presumed nativeness* ($F(1, 6399.1) = 59.29, p = .00$), and *listener group* ($F(3, 4818.5) = 32.64, p = .00$).

Figs. 2 and 3 also show that the *status* evaluations for speakers that were considered L1 or L2 English speakers showed a substantial overlap within and between listener groups. This suggests that regardless of listener group, there were generally not many differences between speakers in terms of the *status* they were given on the basis of presumed nativeness. What is striking is that, for all listener groups, the speakers that were considered to be L2 English speakers, regardless of their actual English accent, had more varied *status* evaluations, resulting in wider confidence intervals compared with speakers that were viewed as L1 English speakers.

There was a two-way interaction between *speaker* and *listener group* ($F(24, 6070.4) = 2.51, p = .00$), meaning that the *status* for each speaker differed significantly per listener group, which was the case for all verbal-guise speakers compared with the matched-guise speaker. The lowest *status* evaluation was given by Dutch listeners to the Dutch English matched-guise speaker when he was believed to be an L2 English speaker, which was in the majority of cases (>96%). However, in the minority of cases where the Dutch English matched-guise speaker was considered an L1 English speaker by Dutch listeners, his *status* evaluations increased significantly, but with a much wider confidence interval (see Figs. 2 and 3 MG_DE, red dot/line is the Dutch listener group). Furthermore, in general, *presumed nativeness* did affect *status* evaluations, with speakers that were viewed as being an L1 English speaker ($M = 3.87, SD = .63$) being attributed higher *status* than speakers who were not ($M = 3.57, SD = .73$) ($F(1, 6340.7) = 46.90, p = .00$).

3.2.2. Affect and presumed nativeness

A linear mixed model analysis (Figs. 4, 5) showed that there was no three-way interaction between *speaker*, *presumed nativeness*, and *listener group* ($F(18, 6351.4) = 1.39, p = .13$), as well as no two-way interaction *presumed nativeness* and *listener group* ($F(3, 6351.0) = 1.29, p = .28$). However, there was a two-way interaction between *speaker* and *presumed nativeness* ($F(6, 6340.7) = 9.67, p = .00$), and between *speaker* and *listener group* ($F(18, 6253.1) = 2.33, p = .001$). This means that *affect* evaluations varied significantly for individual speakers, depending on whether listeners had given them L1 or L2 status, and *affect* evaluations for individual speakers also differed between listener groups, even though listener group (country) in general did not impact *affect* evaluations ($F(3, 5027.5) = 1.31, p = 0.27$). The *affect* evaluations were, however, impacted by *presumed nativeness* ($F(1, 6396.3) = 4.11, p = .04$). Correctly identifying an L1 English speaker (both matched guises and verbal guises) as being an L1 English speaker led to more consistent *affect* evaluations, observed in narrower confidence intervals when compared to speakers given L2 English speaker status (Figs. 4, 5); and *affect* was higher for speakers viewed as L1 English ($M = 3.53, SD = .76$) compared to speakers who were not ($M = 3.38, SD = .74$).

3.2.3. Dynamism and presumed nativeness

A linear mixed model analysis (Figs. 6, 7) showed that there was no three-way interaction between *speaker*, *presumed nativeness*, and *listener group* ($F(18, 6229.1) = 1.47, p = .09$), and no two-way interaction between *speaker* and *presumed nativeness* ($F(6, 6221.3) = 1.51, p = .17$). There was a two-way interaction between *presumed nativeness* and *listener group* ($F(3, 6332.3) = 3.09, p = .03$), between *speaker* and *listener group* ($F(18, 6125.4) = 1.74, p = .03$) and *listener group* significantly impacted *dynamism* evaluations ($F(3, 4806.4) = 12.69, p = .00$). This means that *dynamism* evaluations varied significantly between *listener groups* in general, but also for an individual *speaker* when they were presumed to be an L1 English speaker. For example, Dutch listeners ($M = 3.10, SD = .78$) assigned lower *dynamism* than German listeners ($M = 3.30, SD = .83$), Spanish listeners ($M = 3.44, SD = .79$), and Singaporean ($M = 3.34, SD = .76$) listeners, and German and Singaporean listeners assigned significantly lower *dynamism* than Spanish listeners. In addition, one standard British English speaker (V4 BE) was considered significantly more *dynamic* than all other speakers, but only when he was assigned L1 English speaker status. This might mean that the correct combination of accent, *presumed nativeness*, and an L1 (standard, British) English accent, produced by a speaker with a particular voice quality can lead to significantly more consistent and more positive evaluations of *dynamism*.

Furthermore, *presumed nativeness* impacted *dynamism* ($F(1, 6404.9) = 11.40, p = .001$) in that speakers that were given L1 English speaker status ($M = 3.36, SD = .79$) were viewed as more dynamic than speakers who were not viewed as L1 English speakers ($M = 3.14, SD = .81$). An L1 English accent produced by both a matched-guise and verbal-guise speaker that was viewed as coming from an L1 English speaker status was given more consistent *dynamism* evaluations compared to that same L1 English speaker when they were considered to be an L2 English speaker. Fig. 7 also shows more consistent *dynamism* evaluations for the Dutch English speakers that were viewed as L2 English speakers compared with the L1 English speakers that were viewed as L2 English speakers.

In summary, our analyses clearly show that speaker evaluations are significantly affected by *presumed nativeness*, for all dimensions: *status*, *affect*, *dynamism*. What is interesting is that the *status*, *affect*, and *dynamism* evaluations of

Table 7

Significant effects (positive = +, negative = −, or no effect = 0) of *accent*, *presumed nativeness* and their interaction on *speaker evaluations* (status, affect, dynamism), per speaker group (MG = matched guises, VG = verbal guises) and per *listener group* (NL, Germany, Spain, Singapore). Cell pairs in gray indicate conflicting evidence between MG and VG.

Listener group	Factor	Status		Affect		Dynamism	
Speaker		MG	VG	MG	VG	MG	VG
NL	Accent	+	+	+	0	0	0
	Presumed nativeness	+	0	0	0	0	+
	Accent x Presumed nativeness	0	0	0	0	0	0
Germany	Accent	0	0	0	0	0	0
	Presumed nativeness	0	+	0	0	0	0
	Accent x Presumed nativeness	0	0	0	0	0	0
Spain	Accent	0	0	0	0	0	0
	Presumed nativeness	0	0	0	0	0	0
	Accent x Presumed nativeness	0	0	0	0	0	0
Singapore	Accent	0	0	0	0	0	0
	Presumed nativeness	0	+	0	0	0	0
	Accent x Presumed nativeness	0	0	0	0	0	0

L1 and L2 English speakers, regardless of *presumed nativeness*, did not differ significantly between speakers and listener groups. However, for the Dutch listener group, when the Dutch-accented English speakers were perceived to be L2 English speakers, this led to significantly lower *status* compared to instances where they were identified as L1 English speakers of Dutch. *Affect* and *dynamism* were significantly affected to an equal degree by *presumed nativeness*, since (correctly) assigning nativeness led to the most consistent *affect* and *dynamism* evaluations. This consistency implies that when the listener groups, each of whom have varying language backgrounds, were sure about the speakers' (non-)nativeness, they more easily tapped into a language norm that associates (non-)nativeness with specific degrees of *affect* and *dynamism*. In general, however, the effects of accent were not strong, and there were also general differences observed for individual speakers, for example, the extremely high dynamism that one verbal-guise speaker triggered. This led us to investigate the impact of *voice* and *accent* on *speaker evaluations* in more detail.

3.3. The impact of voice and accent on speaker evaluations

To understand the extent to which *voice* impacts responses to speech samples in greater depth, we compared the *speaker evaluations* of the matched guises with the speaker evaluations of the verbal guises. The statistical analyses

Table 8

Significant effects (positive = +, negative = −, or no effect = 0) of *accent* or *voice*, *presumed nativeness*, *accent* or *voice* and *presumed nativeness* on speaker evaluations (*status*, *affect*, *dynamism*), for the six verbal-guise speakers, per *listener group* (NL, Germany, Spain, Singapore). The predictor is *Voice* (six voices) or *Accent* (three accents, each accent with two voices). The accent outcomes are the same as those in Table 6 for VG (verbal guise). Cell pairs in gray indicate conflicting evidence between status and accent.

Verbal Guises	Factor	Status		Affect		Dynamism	
		Accent	Voice	Accent	Voice	Accent	Voice
NL	Predictor	+	+	0	+	0	+
	Presumed nativeness	0	+	0	0	+	0
	Predictor x Presumed nativeness	0	0	0	0	0	0
Germany	Predictor	0	0	0	+	0	+
	Presumed nativeness	+	+	0	0	0	0
	Predictor x Presumed nativeness	0	0	0	+	0	0
Spain	Predictor	0	0	0	+	0	+
	Presumed nativeness	0	0	0	0	0	0
	Predictor x Presumed nativeness	0	0	0	0	0	0
Singapore	Predictor	0	0	0	0	0	+
	Presumed nativeness	+	0	0	0	0	0
	Predictor x Presumed nativeness	0	0	0	+	0	0

were all conducted with linear mixed models, as reported in the previous sections (see Supplementary Materials 3 via <https://surfdive.surf.nl/files/index.php/s/FFMZCCrttF7xIGC> for means and SDs).

Table 7 shows the effects of *accent* and *presumed nativeness* per *speaker* (matched-guise speaker, verbal-guise speakers) for each *listener group*. The speaker evaluations for *status*, *affect*, and *dynamism* were all impacted by individual speakers. For Dutch listeners, the responses to the accentedness of both the verbal-guise speakers and matched-guise speaker indicated a significantly positive effect of *accent* on *status*, meaning that an L1 English accent, whether it was produced by the matched-guise speaker or the verbal-guise speakers, evoked a higher status compared with Dutch-accented English.

For the matched-guise speaker, *presumed nativeness* also positively influenced *status*, meaning that when Dutch listeners believed the matched-guise speaker to be an L1 English speaker, regardless of the actual accent he produced, this positively influenced *status* evaluations of the matched-guise speaker. Furthermore, when the matched-guise speaker spoke standard British English or Dutch-accented English, *affect* was significantly higher compared to when he spoke standard American English. In terms of *dynamism*, a positive effect of *presumed nativeness* was observed for the verbal guises. There was no combined effect of *accent* and *presumed nativeness* on *speaker evaluations* for either verbal or matched guises.

For German and Singaporean listeners, we saw a positive effect of *presumed nativeness* of the verbal guises on *status*, meaning that verbal guise speakers that were viewed as being L1 English speakers were evaluated as having higher status compared to when the verbal guise speakers were viewed as L2 English speakers. There were no other (combined) effects of *accent* and *presumed nativeness* on *speaker evaluations*. For Spanish listeners, there were no (combined) effects of *accent* and *presumed nativeness* on the *speaker evaluations* of either the matched or verbal guises.

Accent and *presumed nativeness* appeared to have had relatively limited effects on *speaker evaluations*, especially for the German, Spanish and Singaporean listeners; but since we also observed significantly different *speaker evaluations* of individual verbal-guise speakers, we investigated the potential effect of *voice* compared to *accent* on *speaker evaluations* (*status*, *affect*, *dynamism*). Table 8 shows the effects of *accent*, *presumed nativeness* and *voice* on the *speaker evaluations* of all six verbal-guise speakers.

For Dutch listeners, we can see that not only *accent*, but also *voice*, affected *status evaluations* positively, with the Dutch-accented English verbal-guise speaker 1 (V1) being perceived as having a significantly higher status than the other (V1). This clearly shows how an individual's voice can impact *status evaluations*. In addition, *affect* and *dynamism* were also positively affected by *voice*, with the British and American verbal-guise speakers differing significantly from one another in terms of the *affect* and *dynamism* they evoked.

German and Spanish listeners displayed similar patterns to Dutch listeners in terms of *voice*, which had a significantly positive impact on *affect* and *dynamism* evaluations only. For Singaporean listeners, the pattern was similar but only for *dynamism*. As with the Dutch listeners, the other three listener groups also seemed to prefer the same verbal-guise speaker over the other within one accent (e.g. V3,4 for standard British English; V5,6 for standard American English). German listeners showed a significantly positive relationship between *presumed nativeness* and *voice* for *status*, which meant that when a speaker was considered to be an L1 English speaker and had a voice quality that we assume most likely appealed to the listener, the *status* of this speaker was positively impacted (e.g. V3 for standard British English). Finally, Singaporean listeners only showed a preference for a speaker in relation to *affect*, which was positively impacted when an individual speaker was viewed as both an L1 English speaker (*presumed nativeness*) and had a voice quality that most likely appealed to the listener.

4. DISCUSSION AND CONCLUSION

In this section we discuss the answers to our research questions, our study's limitations, and implications for future accentedness research.

4.1. RQ1 *presumed nativeness* in an ELF context and its impact on *speaker evaluations*

Firstly, we wanted to determine the extent to which L2 English listener groups associate the tested accents with standard British and American English accents and Dutch-accented English (non-)nativeness, and whether associating a speaker's accent with (non-)nativeness impacts *speaker evaluations*.

What our results indicate is that, in general, all four listener groups (Dutch, German, Spanish, Singaporean) correctly associated the two standard English accents with L1 English speech, and especially the standard British English accent, even though they did not always know which individual L1 English accent they had listened to. These results suggest that, as with L1 (English) listeners, both L1 and L2 English speakers can easily distinguish between L1 and L2 English accent varieties, but not necessarily between L1 English accent varieties (Brunner, 2009; Gnevshева, 2018; McGorrey and McMahon, 2017; Nolan, 2003; Wong and Babel, 2017). This may be because while L1 English accents are significantly similar phonetically, they are phonetically significantly different from L2 English accents (see also McMahon et al., 2007). These results also suggest that our L2 English listeners were likely to be familiar with at least one of the L1 English accent varieties. Indeed, the best-recognized accent variety in our study was British English, which is the variety traditionally used in general education and commercial language training in Europe (the Netherlands, Germany, Spain) and which is also historically connected to Singapore and still used in Singaporean teacher training (British Council, 2024).

In terms of identifying the L2 English accent, it is striking that in instances where the Dutch-accented English speech samples were considered to be L2 English, they were generally viewed by all listener groups as having European origins, even though the content of the speech samples in no way alluded to a European setting or included European references. It might be the case that, to the listeners, the Dutch-accented English samples possessed phonetic and prosodic features that suggested European origins, and as a result the listeners selected a European country or just "Europe" as the speaker's origin.

Interestingly, 9.0–16.2% of German, Spanish and Singaporean listeners believed that the accent origin of one Dutch-accented English verbal-guise speaker (V2) was Indian. More specifically, some Singaporeans indicated that the Dutch-accented English accent matched and verbal guises (5.4–12.4%) were of Indian origin. This might be because the Singaporean listeners who believed that the speech samples were L2 English accents compared them to the L2 English accents they were themselves familiar with. Singapore is a multi-ethnic society whose residents are mostly of Chinese descent, followed by Singaporeans of Malay or Indian descent (less than 10%) on the basis of paternal data (Department of Statistics Singapore, 2023). The Singaporean listeners might have perceived the

Dutch-accented English accent as phonetically and prosodically most similar to an Indian-accented English, with which they could be relatively familiar due to the makeup of Singaporean society, and concluded that this was the origin of the Dutch-accented English accent.

It is worth noting that there were a few striking differences in accent identification patterns between listener groups. For instance, the Dutch-accented English of one verbal-guise speaker (V1 DE) was mostly considered L1 English by three of the L2 English listener groups (German, Singaporean, Spanish), who, with the exception of German listeners, were unlikely to be familiar with Dutch or Dutch-accented English. For the matched guise and other verbal guises, approximately a third of these three listener groups indicated that the Dutch-accented English speakers were in fact L1, in most cases British English speakers. In contrast, the Dutch listeners displayed their familiarity with Dutch-accented English as they were best at indicating the origins of the Dutch-accented English speech stimuli, with a minimum of 75% of Dutch listeners correctly identifying both the Dutch-accented English matched and verbal guises.

In terms of presuming a speaker's (non-)nativeness and its impact on speaker evaluations, we conclude that (correctly) perceiving someone as an L1 English speaker leads to more consistent speaker evaluations (*status*, *affect*, *dynamism*) compared to perceiving someone as an L2 English speaker, even if the specific L1 English accent is not (correctly) identified. In addition, being viewed as an L1 English speaker seems to influence speaker evaluations positively across all dimensions (*status*, *affect*, and *dynamism*), although in general the effects of accent were not strong. Interestingly, a speaker's individual *status* varied per listener group. For example, there were differences between Dutch listeners and their German, Spanish and Singaporean counterparts, with the Dutch assigning significantly lower *status* when they correctly identified Dutch-accented English as L2 English and as coming from an L1 speaker of Dutch. It appears that for L2 English listeners, as with L1 English listeners (see also Nejari et al., 2012), sufficient familiarity with a specific L2 English accent can lead to significantly lower *status*. It might be the case that in the minds of Dutch listeners, Dutch-accented English in particular carries a low social status compared with L1 English: this might be because they have very specific associations, based on experience and exposure, or even experience vicarious shame (see also Hendriks et al., 2016), with Dutch-accented English and its speakers. It is therefore conceivable that a similarly strong response to Dutch listeners' "own" English accent variety might not be observed for other L2 English listeners and "their" English accent varieties. Dissimilar *status*, *affect* and *dynamism* evaluations were more consistent for Dutch-accented English speakers identified as L2 English speakers than for Dutch-accented English speakers considered to be L1 English. This was the case for all four listener groups, and suggests that the listeners were more confident, and hence more consistent, in their speaker evaluations when they correctly identified the Dutch-accented English as being L2 in origin.

What the results in general suggest is that an essential part of speaker evaluations (in an ELF professional communication context) is the estimation of a speaker's nativeness, and that, when this estimation is correct, speaker evaluations are more consistent. Furthermore, listeners might be most confident applying language norms, and evaluating speakers on the basis of this language norm, when they know 'who' they are listening to. The language norm that appears to be applied by L2 English listeners is one that associates nativeness (as opposed to non-nativeness) with more positive characteristics in terms of social status, *affect* (or likeability), and *dynamism* (or a person's enthusiasm and proactivity). As a result, earlier conclusions with regard to the lack of relevance of accentedness in perceptions of speakers in an ELF context might not be complete (Canagarajah, 2007; Nejari et al., 2021). It is not so much that accentedness is unimportant to L2 English speakers in ELF contexts, but more that L2 English speakers may tend mistakenly to categorize and presume the origins of English accents: this can then make it seem as if their responses reflect a non-traditional language norm, when this is not the case.

4.2. RQ2. Speaker voice or accent impacting speaker evaluations?

On the basis of responses to individual speakers, we wanted to determine whether a speaker's voice impacted speaker evaluations. We therefore assessed whether listeners responded differently to matched guises when compared to verbal guises, in terms of *accent*, *presumed nativeness* and *speaker evaluations* (*status*, *affect*, *dynamism*). This also allowed us to compare the effectiveness of matched guises and verbal guises in accentedness research that focused on L1 and L2 English accents and to investigate the extent to which a speaker's voice impacts listeners' responses. We found that, in the majority of cases, matched guises and verbal guises lead to largely similar effects, illustrating that both can be used to effectively assess *presumed nativeness* and *speaker evaluations*.

When we further analyzed the responses to individual speakers, we observed some interesting patterns. For Dutch listeners, we found that a speaker's voice affected *status*, with Dutch English verbal-guise speaker 1 (V1) being perceived as having a significantly higher *status* than Dutch English verbal-guise speaker 2 (V2). It should be noted that Dutch listeners did so, whilst in the majority of cases correctly identifying the Dutch-accented English speech samples and assigning them lower *status*. Hence, despite knowing the language background of V1, the Dutch listeners assigned him a significantly higher status compared to V2, demonstrating how *voice* is most likely to have impacted his *status*.

evaluations. In addition, *affect* and *dynamism* were also positively affected by *voice*, with almost all verbal-guise speakers within one accent-type differing significantly from one another in terms of the *affect* and *dynamism* they evoked, and with similar accent identification patterns (e.g., V3 lower *dynamism* than V4 for standard British English; V5 lower *affect* than V6 for standard American English).

In terms of *affect* and *dynamism*, we can again observe the impact of a speaker's voice in the preference of most listener groups for one verbal-guise speaker of a particular accent over the other verbal guise speaker with that same accent, with respect to the tested L1 English accents. Future research might investigate whether similar patterns would emerge for a wide range of L1 and L2 English listener groups and which voice(s) qualities evoke positive speaker evaluations among specific groups.

The varying responses to the verbal guise speakers' voices are interesting in the context of general speaker evaluation processes. Our results appear to suggest that these processes are based on estimations of a speaker's language background and that perceived nativeness functions are an important prerequisite for speaker evaluations to be triggered. However, speaker evaluations also appear to be influenced by an assessment of a speaker's voice quality, and if an individual's voice quality happens to stand out positively or negatively to listeners, the subsequent outcome may be that a speaker is evaluated more positively or negatively, despite their (non-)nativeness. This confirms that voice, like accent, is a fundamental aspect of speaker evaluations (see also Doeleman, 1998).

4.3. Limitations, implications, and future research

In our study, we tested two L1 English varieties and one L2 English variety from a language that is part of the same language family as English, namely the West-Germanic. Whereas our listeners were not able to identify the origin of the Dutch-accented English speech samples, they did mostly assign them European, and often even L1 English, origins. A limitation of our study is that we did not collect information on the familiarity of the listeners with specific languages and varieties, or assess how confident listeners felt about their answers on the country of origin of the speakers. We assumed that the German, and even more so, Spanish and Singaporean listeners, were unfamiliar with Dutch and Dutch-accented English; but this was something that we did not ask. Even with the German listeners, who were the best at identifying Dutch-accented English (40%) of the three listener groups, we did not assume their familiarity: this was because the experiment was conducted across Germany as a whole, not necessarily in those areas bordering the Netherlands.

More knowledge on presumed (non-)nativeness can give us a better understanding of the role familiarity plays and when an accent variety is considered L1 or not: for instance, on the basis of specific phonetic features. As a result, we can better understand whether accent varieties from certain language families are also presumed to originate from specific language families and regions in the world, when the accent is heard in another language, in our case English. Future research should therefore focus on multiple L2 English accents, their phonetic and prosodic features, and how these are categorized and evaluated by a variety of L2 English listeners. While the current study used speakers that were regarded as having a natural and authentic accent, who were used to speaking in public in a professional capacity (see also Nejari et al., 2019), and comparable in terms of their voice age, we did not control for other potential voice differences in aspects such as pitch and speech tempo. Future studies could investigate a broader age range of speakers (displaying potentially greater differences in voice characteristics) to determine the role such factors play in the process of speaker evaluations.

In terms of employing the matched-guise and verbal-guise technique in an ELF accentedness study, we can confidently state that matched guises offer a valid method to investigate how listeners assess L1 and L2 English speakers. Comparing responses to both matched and verbal guises has illustrated the strength of matched guises when the objective in a study is to eliminate the impact of a speaker's voice. We observed consistent responses to the matched guises and observed significant differences in the responses to the three accents produced by the same speaker, and hence the same voice. In contrast, we observed some striking speaker evaluations of certain individual verbal-guise speakers, regardless of accent, for example when it comes to *dynamism*. One British English verbal-guise speaker evoked significantly higher dynamism, when compared to the other British English verbal-guise speaker, but also in comparison to all the other speakers. In terms of applying the matched-guise technique to assess responses to both L1 and L2 English accents, it will remain challenging to find matched-guise speakers who can representatively and authentically produce a combination of L1 and L2 English accents, let alone ones from several language families (as discussed in Nejari et al., 2019). Verbal guises then offer the most practical alternative, even though our results highlight the potential issue with the verbal-guise technique: a speaker's voice and prosodic factors can have a significant influence on listeners' evaluations.

Since research has shown that voice use varies per culture, context, and even individual, it might be the case that preferences for specific voices, for instance in terms of intonation and tone, can also be dependent on a listener's cul-

tural or even linguistic background (e.g., Majid, 2012). Future research could make a special focus on this question to establish which accents and voices, separately and together, create the best impression with others. This in turn might be of use in language learning or audio technology designed for commercial or informational purposes. Finally, the language norm applied by fluent L2 speakers of English in this study illustrates that even though comprehension does not have to be hindered by unfamiliar accents, traditional perspectives on the superiority of nativeness would still seem to apply, even in an ELF context in which L2 English varieties are the standard. Foreign language education therefore should focus on developing more tolerance to the full range of L2 English varieties to ensure unbiased evaluations of the abilities and character of L2 English speaker groups.

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Warda Nejari: Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **Roeland van Hout:** Writing – review & editing, Visualization, Methodology, Formal analysis. **Marinel Gerritsen:** Writing – review & editing. **Brigitte Planken:** Writing – review & editing.

Data availability

Data will be made available on request.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Bond, C.F., DePaulo, B.M., 2006. Accuracy of Deception Judgments. *Pers. Soc. Psychol. Rev.* 10 (3), 214–234. https://doi.org/10.1207/s15327957pspr1003_2.
- Bresnahan, M.J. et al., 2002. Attitudinal and affective response toward accented English. *Lang. Commun.* 22 (2), 171–185. [https://doi.org/10.1016/S0271-5309\(01\)00025-8](https://doi.org/10.1016/S0271-5309(01)00025-8).
- Brunner, E.G., 2009. The study of variatio from two perspectives. *Lang. Linguist. Compass* 3 (3), 734–750. <https://doi.org/10.1111/j.1749-818X.2009.00137.x>.
- Buckingham, L., 2014. Attitudes to English teachers' accents in the Gulf. *Int. J. Appl. Linguist.* 24, 50–73.
- Canagarajah, S., 2007. Lingua Franca English, multilingual communities, and language acquisition. *Mod. Lang. J.* 91 (1), 923–939. <https://doi.org/10.1111/j.1540-4781.2007.00678.x>.
- Cargile, A.C., Giles, H., 1997. Understanding language attitudes: exploring listener affect and identity. *Lang. Commun.* 17 (3), 195–217. [https://doi.org/10.1016/S0271-5309\(97\)00016-5](https://doi.org/10.1016/S0271-5309(97)00016-5).
- Carlson, H.K., McHenry, M.A., 2006. Effect of accent and dialect on employability. *J. Employ. Couns.* 43 (2), 70–83. <https://doi.org/10.1002/j.2161-1920.2006.tb00008.x>.
- British Council, 2024. Why learn English with us? <https://www.britishcouncil.sg/english-courses/why-learn>
- Cowen, A.S. et al., 2019. The primacy of categories in the recognition of 12 emotions in speech prosody across two cultures. *Nat. Hum. Behav.* 3, 369–382. <https://doi.org/10.1038/s41562-019-0533-6>.
- Da Silva, C.S., Leach, A.-M., 2013. Detecting deception in second-language speakers. *Leg. Criminol. Psychol.* 18 (1), 115–127. <https://doi.org/10.1111/j.2044-8333.2011.02030.x>.
- De Lange, F.P., Heilbron, M., Kok, P., 2018. How do expectations shape perception? *Trends Cogn. Sci.* 22 (9), 764–779. <https://doi.org/10.1016/j.tics.2018.06.002>.
- Department of Statistics Singapore, 2023. Resident population by ethnic group. <https://www.singstat.gov.sg/find-data/search-by-theme/population/population-and-population-structure/visualising-data/resident-population-by-ethnic-group-age-group-and-sex-dashboard>
- Ding, H., Hoffmann, R., Jokisch, O., 2018. Voice preferences in German: a cross-linguistic comparison of native and Chinese listeners. *Proceedings of the 29th Conference on Electronic Speech Signal Processing*. https://www.essv.de/pdf/2018_105_112.pdf
- Doeleman, R., 1998. *Native Reactions to Nonnative Speech*. Tilburg University Press.
- Dragojevic, M., Berglund, C., Blauvelt, T.K., 2018. Figuring out who's who: the role of social categorization in the language attitudes process. *J. Lang. Soc. Psychol.* 37 (1), 28–50. <https://doi.org/10.1177/0261927X17706942>.
- Eichhorn, J.T. et al., 2018. Effects of aging on vocal fundamental frequency and vowel formants in men and women. *J. Voice* 32 (5), 644–e1. <https://doi.org/10.1016/j.jvoice.2017.08.003>.

- Fuertes, J.N. et al., 2012. A meta-analysis of the effects of speakers' accents on interpersonal evaluations. *Eur. J. Soc. Psychol.* 42 (1), 120–133. <https://doi.org/10.1002/ejsp.862>.
- Garrett, P., 2010. *Attitudes to language. Key Topics in Sociolinguistics.* Cambridge University Press.
- Gnevnsheva, K., 2018. Variation in foreign accent identification. *J. Multiling. Multicult. Dev.* 39 (8), 688–702. <https://doi.org/10.1080/01434632.2018.1427756>.
- Grondelaers, S., Van Hout, R., 2015. How (in)coherent can standard languages be? A perceptual perspective on co-variation. *Lingua* 172–173, 62–71. <https://doi.org/10.1016/j.lingua.2015.10.012>.
- Hendriks, H., 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 J. Appl. Linguist.* 5 (1), 1–17. <https://doi.org/10.1075/dujal.5.1.01hen>.
- Hosoda, M., Nguyen, L.T., Stone-Romero, E.F., 2012. The effect of Hispanic accents on employment decisions. *J. Manag. Psychol.* 27 (4), 347–364. <https://doi.org/10.1108/02683941211220162>.
- Jenkins, J., 2009. English as a Lingua Franca: interpretations and attitudes. *World Englishes* 28 (2), 200–207. <https://doi.org/10.1111/j.1467-971X.2009.01582.x>.
- Kajjura, M. et al., 2021. Brain activity predicts future learning success in intensive second language listening training. *Brain Lang.* 212. <https://doi.org/10.1016/j.bandl.2020.104839>.
- Kalin, R., Rayko, D., Love, N., 1980. The perception and evaluation of job candidates with four different ethnic accents. In: Giles, H., Robinson, W.P., Smith, P. (Eds.), *Language: Social Psychological Perspectives.* Pergamon, pp. 197–202. <https://doi.org/10.1016/B978-0-08-024696-3.50034-X>.
- Kreiman, J. et al., 1993. Perceptual evaluation of voice quality: review, tutorial, and a framework for future research. *J. Speech Hear. Res.* 36 (1), 21–40. <https://doi.org/10.1044/jshr.3601.21>.
- Leach, A.M., Da Silva, C.S., 2013. Language proficiency and police officers' lie detection performance. *J. Police Crim. Psychol.* 28, 48–53. <https://doi.org/10.1007/s11896-012-9109-3>.
- Levy-Ari, S., Keysar, B., 2010. Why don't we believe non-native speakers? The influence of accent on credibility. *J. Exp. Soc. Psychol.* 46 (6), 1093–1096. <https://doi.org/10.1016/j.jesp.2010.05.025>.
- Lindemann, S., 2003. Koreans, Chinese, or Indians? Attitudes and ideologies about nonnative English speakers in the United States. *J. Socioling.* 7 (3), 348–364. <https://doi.org/10.1111/1467-9481.00228>.
- Lortie, C. et al., 2018. Age differences in voice evaluation: from auditory-perceptual evaluation to social interactions. *J. Speech Lang Hear. Res.* 61 (1). https://doi.org/10.1044/2017_JSLHR-S-16-0202.
- Majid, A., 2012. Current emotion research in the language sciences. *Emot. Rev.* 4 (4), 432–443. <https://doi.org/10.1177/1754073912445827>.
- McBride, K., 2011. The effect of rate of speech and distributed practice on the development of listening comprehension. *Comput. Assist. Lang. Learn.* 24 (2), 131–154. <https://doi.org/10.1080/09588221.2010.528777>.
- McGorrey, P.G., McMahon, M., 2017. A fair 'hearing': earwitness identifications and voice identification parades. *Int. J. Evid. Proof* 21 (3), 262–286. <https://doi.org/10.1177/1365712717690753>.
- McKenzie, R.M., 2008. Social factors and non-native attitudes towards varieties of spoken English: a Japanese case study. *Int. J. Appl. Linguist.* 18 (1), 63–88. <https://doi.org/10.1111/j.1473-4192.2008.00179.x>.
- McMahon, A. et al., 2007. The sound patterns of Englishes: representing phonetic similarity. *English Lang. Linguist.* 11 (1), 113–142. <https://doi.org/10.1017/S1360674306002139>.
- Miyamoto, K. et al., 2021. Identification and disruption of a neural mechanism for accumulating prospective metacognitive information prior to decision-making. *Neuron* 109 (8), 1396–1408. <https://doi.org/10.1016/j.neuron.2021.02.024>.
- Munro, M.J., Derwing, T.M., 1995a. Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Lang. Learn.* 45 (1), 73–97. <https://doi.org/10.1111/j.1467-1770.1995.tb00963.x>.
- Munro, M.J., Derwing, T.M., 1995b. Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Lang. Speech* 38 (3), 289–306. <https://doi.org/10.1177/002383099503800305>.
- Nejjari, W. et al., 2012. Responses to Dutch-accented English. *World Englishes* 31 (2), 248–267. <https://doi.org/10.1111/j.1467-971X.2012.01754.x>.
- Nejjari, W. et al., 2019. Refinement of the matched-guise technique for the study of the effect of non-native accents compared to native accents. *Lingua* 219, 90–105. <https://doi.org/10.1016/j.lingua.2018.12.001>.
- Nejjari, W. et al., 2020. Where does a 'foreign' accent matter? German, Spanish and Singaporean listeners' reactions to Dutch-accented English, and standard British and American English accents. *PLoS One* 15 (4), e0231089. <https://doi.org/10.1371/journal.pone.0231089>.
- Nejjari, W. et al., 2021. Dutch listeners' responses to Dutch, British and American English accents in three contexts. *Dutch J. Appl. Linguist.* 10. <https://doi.org/10.51751/dujal9365>.
- Nolan, F., 2003. A recent voice parade. *Int. J. Speech, Lang. Law* 10 (2), 277–291. <https://doi.org/10.1558/sll.2003.10.2.277>.
- Pajupuu, H., Altrov, R., Pajupuu, J., 2019. The effects of culture on voice likability. *Trames J. Hum. Soc. Sci.* 23, 239–257. <https://doi.org/10.3176/tr.2019.2.08>.
- Pell, M.D. et al., 2009. Recognizing emotions in a foreign language. *J. Nonverbal Behav.* 33, 107–120. <https://doi.org/10.1007/s10919-008-0065-7>.
- Perry, L.K. et al., 2018. Influences of speech familiarity on immediate perception and final comprehension. *Psychon. Bull. Rev.* 25, 431–439. <https://doi.org/10.3758/s13423-017-1297-5>.

- Pietraszewski, D., Schwartz, A., 2014. Evidence that accent is a dimension of social categorization, not a byproduct of perceptual salience, familiarity, or ease-of-processing. *Evol. Hum. Behav.* 35 (1), 43–50. <https://doi.org/10.1016/j.evolhumbehav.2013.09.006>.
- 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 Jyväskylä, Finland, 1997). *Diss. Abstr. Int.* 60, 9C.
- Rojas, S., Kefalianos, E., Vogel, A., 2020. How does our voice change as we age? A systematic review and meta-analysis of acoustic and perceptual voice data from healthy adults over 50 years of age. *J. Speech Lang. Hear. Res.* 63 (2), 533–551. https://doi.org/10.1044/2019_JSLHR-19-00099.
- Van Engen, K.J., Peelle, J.E., 2014. Listening effort and accented speech. *Front. Hum. Neurosci.* 8. <https://doi.org/10.3389/fnhum.2014.00577>.
- Wong, P., Babel, M., 2017. Perceptual identification of talker ethnicity in Vancouver English. *J. Socioling.* 21, 603–628. <https://doi.org/10.1111/josl.12264>.
- Yiu, E.M. et al., 2008. Cultural and language differences in voice quality perception: a preliminary investigation using synthesized signals. *Folia. Phoniater. Logop.* 60 (3), 107–119. <https://doi.org/10.1159/000119746>.
- Yook, C., Lindemann, S., 2013. The role of speaker identification in Korean university students' attitudes towards five varieties of English. *J. Multiling. Multicult. Dev.* 34 (3), 279–296. <https://doi.org/10.1080/01434632.2012.734509>.
- Zahn, C.J., Hopper, R., 1985. Measuring language attitudes: the speech evaluation instrument. *J. Lang. Soc. Psychol.* 4 (2), 113–123. <https://doi.org/10.1177/0261927X8500400203>.
- Zoghbor, W.S., 2018. Teaching english pronunciation to multi-dialect first language learners: the revival of the Lingua Franca Core (LFC). *System* 78, 1–14. <https://doi.org/10.1016/j.system.2018.06.008>.