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

Warda Nejjari¹

¹Affiliation not available

December 13, 2022

Abstract

We investigated to what extent responses (N=6617) by four non-native English listener groups (The Netherlands: n=1701; Germany: n=1606; Spain: n=1647; Singapore: n=1663) assigned nativeness to standard, native British and American English accents, compared with a typical Dutch English accent. We assessed to what extent assigned nativeness impacted speaker evaluations (status, affect, dynamism), and to what extent a speaker’s voice influenced speaker evaluations by analyzing listener responses to verbal and matched guises. Results showed that perceptions of a speaker’s nativeness significantly impacted speaker evaluations on all dimensions, and therefore we conclude that speaker evaluations are also based on perceptions of the nativeness of a speaker. 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 native and non-native English speakers. Finally, this study confirms the relevance and superiority of the matched-guise technique in accentedness research, compared with the verbal-guise technique, since the former successfully minimizes the actual impact of voice.

Introduction

In general, speaker evaluation research focusing on non-native English accents shows that (strong) non-native English accents, compared to native English accents, evoke more negative speaker evaluations in native English listeners, in terms of social status and (perceived) speech comprehension, but not necessarily regarding the affect a speaker evokes or a speaker’s dynamism (e.g., Cargile & Giles 1997, Fuertes et al. 2012, Hendriks et al. 2016, Kalin et al. 1980, Levy-Ari & Keysar 2010, Lindemann 2003, Munro & Derwing 1995a,b, Phiko 1997, Nejjari et al. 2012, 2021). In addition, it has been shown that negative perceptions of non-native English speaker groups by native English speaker groups may even negatively impact perceptions of non-native speakers’ trustworthiness, chances for employment and promotion opportunities in organizations (e.g., Da Silva & Leach, 2013; Leach & Da Silva, 2013; Bond & DePaulo, 2006; Carlson & McHenry 2006; Hosoda et al. 2012). Native English in the context of this article refers to the accents produced by inner circle speakers of English, such as British and American English speakers (Kachru, 1985), and non-native English listeners refer to both expanding circle English speakers (e.g., Dutch, German, and Spanish listeners in Nejjari et al. 2020; 2021), but can also include outer circle speakers (e.g., Singaporean listeners in Nejjari et al. 2020).

Non-native evaluations of non-native speakers

Even though there is now more research oriented towards non-native English listeners evaluating non-native English accents, it is still relatively scarce, but provides interesting contrasts to the abovementioned studies. For example, Nejjari et al. (2020) studied how three non-native English listener groups (German, Spanish, Singaporean) perceived Dutch-accented English compared to standard British and American English accents. This study showed that to non-native English listeners who are not familiar with a selected non-native
English accent, the selected non-native English accent does not hinder speech understanding or affect status evaluations, and can evoke similar or even more positive affective evaluations, compared with native, standard English accents. Indeed, other research also shows that when it comes to speech comprehension, unfamiliar, non-native accents can be equally comprehensible as native accents to both (non-)native listeners (see e.g., Perry, Mech & MacDonald 2018; McBride 2011; Van Engen & Peelle, 2014). However, the abovementioned research is focused 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). Therefore, it is reasonable to assume that non-native English accents used in formal, professional communication settings by fluent non-native English speakers can be understood by other fluent non-native English speakers who are not familiar with the used non-native English accent. Also, non-native English speakers do not appear to apply a language norm that connects non-native English accents to lower status, affect and dynamism, compared to native English accents (see also Canagarajah, 2007).

In a follow-up study by Nejjari et al. (2021), in which a Dutch listener group that was familiar with Dutch-accented English reacted to the same three accents, the Dutch listeners, in contrast to the German, Spanish, and Singaporean listeners, evaluated the status of Dutch-accented speech more negatively compared to native, standard English. This suggests that when non-native listeners are familiar with a particular non-native English accent, they readily assign the accent non-native status, and apply a language norm that holds that non-native accented speech reflects a speaker that has lower status, compared to native-accented speech.

Nejjari et al.’s (2020, 2021) opposing results suggest that speaker evaluation processes are at least partly based on assigning (non-)nativeness status to speakers, which is supported by other research. For instance, Yook and Lindemann (2013), studied Korean non-native English listeners’ ability to identify (non-)native English accents and how their awareness of the origins of the accents affected their accent evaluations. Generally, the listeners preferred standard, native-accented English. However, when listeners then had to respond to (non-)native English accents, only listeners who had been made aware of the origins of the tested accents displayed speaker evaluations that followed their previously indicated preference for a native, standard English accent. The listeners who had not been made aware of the accent origins of the speakers did not display these speaker evaluation patterns, which also suggests that, in general, identifying specific accents is challenging to listeners. This is supported by research that shows that, beyond general distinctions between (non-)native accents, or a standard or non-standard accent, listeners often struggle to correctly identify the exact accent they are listening too, unless they are familiar with a particular accent (Gnevsheva 2018; Nolan 2003; Brunner 2009; McGorrery & McMahon 2017; Nejjari et al. 2012; 2020; 2021; Wong & 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, Heilbron, Kok 2018; Miyamoto et al. 2021), also when dealing with language and its speakers (see e.g., Kajiura et al. 2021). This perspective then also suggests that when we listen to someone we apply our available language knowledge and proficiency to predict what a speaker will say, but also, as evolutionary psychology research suggests, use accent to predict which language community a speaker is from, for example, from a native speaker language community, or not (see e.g., Pietraszewski & Schwartz 2014).

In the context of English as a global lingua franca one can imagine that the precise accent origins of non-native English speakers is often not known when interactions between non-native English speakers take place. It is therefore important to know how (non-)native English accent origins are perceived by non-native English listeners, which allows us to better understand how accent labeling impacts speaker evaluations. If it is the case that assigning (non-)nativeness to speakers of English does not impact speaker evaluations, this might mean that for non-native English speakers sounding native-like is not relevant and that English language training should not focus on making learners sound as native-like as possible.

**Voice**

Furthermore, to help establish the precise role that assigning (non-)nativeness plays in speaker evaluation
processes, it is also important to understand whether other speech characteristics, such as a speaker’s voice, influences speaker evaluations. A speaker’s voice communicates important indexical information, such as a speaker’s gender, age, and emotional state with voice having the ability to express approximately 24 emotions (e.g., Majid, 2012), regardless of the language used (Cowen, Elfenbein, Laukka & Keltner 2019; Pell, Monetta, Paulmann & Kotz 2009; Kreiman, Gerratt, Kempster, Erman & Berke 1993). Moreover, an individual’s voice can vary on the basis of cultural background and communication context, but voice use also changes on the basis of relationships people have with specific people and groups. For example, intimate relationships lead people to use their voice differently compared with formal, professional relationships (Ding et al. 2017, 2018; Pajupuu Altrov & Pajupuu 2019; Yiu, Murdoch, Hird, Lau & Ho 2008).

The impact of a speaker’s voice on speaker evaluations can be investigated by conducting speaker evaluation experiments that include both verbal and matched guises. Verbal guises are audio stimuli produced by native speakers of the selected accents, and matched guises are audio stimuli in various selected accents produced by one speaker. Using matched guises is believed to better ensure than verbal guises 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 McGorrery & McMahon 2017 for an overview).

In order to understand the impact of assigning (non-)nativeness status to speakers, but also to what extent voice affects speaker evaluations in the context of English as a lingua franca, we used data collected by Nejjari et al.’s (2020, 2021) in the Netherlands, Germany, Spain, and Singapore to conduct an cross-national, comparative study. Nejjari et al. (2021, 2020) collected data on reactions to matched and verbal guises of native, 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. Native English in the context of this study thus refers to the accents produced by inner circle (British and American) speakers of English (Kachru, 1985) whereas the non-native English listeners in this study were expanding circle English speakers (Dutch, Spanish and German), but also outer circle speakers (Singaporean).

Both the verbal guise and matched guise responses were used as stimuli to investigate to what extent non-native English listener groups assign (non-)native English accent varieties (non-)nativeness status, and how this affects speaker evaluations. Furthermore, we included analyses of reactions to individual speaker’s voices to better understand to what extent speaker evaluation processes are also impacted by a speaker’s voice. This makes it possible to address the following research questions:

To what extent do non-native English listener groups assign (non-)nativeness to standard British and American English accents and Dutch-accented English, and to what extent does assigning (non-)nativeness to speakers impact speaker evaluations?

To what extent does a speaker’s voice impact non-native 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) via an online questionnaire. The experiment had a within-subject multi-factorial design. All listeners (listener groups) were exposed to the independent variables (accent) and evaluated the stimuli on the dependent variables (speaker evaluation, assigned nativeness).

2.1 Speakers

The responses to eight speakers of English were included in the current study. One male matched-guise speaker produced 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 native speakers of Dutch (see Nejjari et al. 2019, 2020, & 2021). The standard accents of British
and American English refer in this study to accents 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 native speakers of Dutch and others familiar with Dutch and Dutch English will recognize as such (see Nejjari et al. 2019, 2020 for further description of phonological features of Dutch-accented English).

We included verbal-guise speech samples by six, male speakers: two native speakers of standard British English, two native speakers of standard American English, and two native speakers of Dutch with a typical Dutch accent in 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 for listeners to compare the other speech samples to, and to see whether the listener responses were consistent. All speakers were aged 35-60 at the time of recording, had at least a master’s degree, and were English language and/or linguistics specialists in some capacity (see Table 1). All but verbal guise speaker six (VG6) had been selected on the basis of a speech sample experiment conducted by Nejjari et al. (2019) that employed intraclass correlation coefficients on reactions by listener groups in The Netherlands, the United States, and Great Britain to determine speech samples that were representative for the three accents. Verbal-guise speaker six had been informally assessed by experienced linguistics professors to represent a typical Dutch-English accent and reactions to this speaker was also included in the current experiment.

Table 1 Speakers per accent

<table>
<thead>
<tr>
<th>Speech samples per accent</th>
<th>One matched-guise speaker</th>
<th>Six verbal guise speakers</th>
<th>One filler speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard British English</td>
<td>MG BE</td>
<td>VG1, VG2</td>
<td>Filler</td>
</tr>
<tr>
<td>Standard American English</td>
<td>MG AE</td>
<td>VG3, VG4</td>
<td></td>
</tr>
<tr>
<td>Dutch English</td>
<td>MG DE</td>
<td>VG5, VG6</td>
<td></td>
</tr>
</tbody>
</table>

2.2 Instrumentation

Stimuli developed by Nejjari et al. (2020; 2021) were used in the current study. One filler text (on a general topic), and three 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 texts and speech sample links in Supplementary Materials). Although communication context was a factor in our earlier studies, 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 on a general topic in standard British English. The speech samples were between 40 and 60 seconds 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 native speakers. To ensure that the speech samples could be evaluated for each accent and speaker, to avoid repeating the content of each context, to limit any order effect, 18 listener groups were created, with a targeted 30 listeners per listener group in each country. This resulted in at least 540 listeners per listener group (Table 2).

On the first page of the questionnaire, the listeners were provided with a general introduction to the questionnaire, 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 and did so by clicking on a link to the speech sample being evaluated, listening to the speech sample, and then answering the question. The listeners were then asked to indicate where they believed a
speaker was from. The answers to the country of origin question were used to determine assigned nativeness. For example, a speaker that was viewed as coming from the United Kingdom was categorized as a speaker that was assigned native English speaker status (assigned nativeness). The listeners also answered comprehension questions for all speech samples. All comprehension or speech understandability scores were high, with little variation between listener groups, so did not impact speaker evaluations (see Nejjari et al. 2021, 2020 for speech understandability results). Finally, listeners were asked to answer questions regarding their English language skills, what native languages they spoke, and biographical details.

2.3. Data collection

The data were collected almost exclusively online in 2016, 2017 and 2018 via Qualtrics, a global survey software and online data collection company that caters for (non-)commercial organizations. The sample of listeners consisted of native speakers of the main national languages of The Netherlands (Dutch), Germany (German), Spain (Spanish), and Singapore (Singapore English) and were highly educated (i.e. had at least reached or completed undergraduate level education). In the Netherlands, the data were collected online via a Facebook page and via Qualtrics. Of the German listener data, 5% was initially collected in the context of a research course at Radboud University in the Netherlands; the remaining 95% was collected via Qualtrics. All the listener data from Spain and Singapore was collected by Qualtrics.

Data collection took place in three rounds for all countries. With each round, incomplete questionnaires (approximately 25%); nonsense answers or symbols and /or only neutral answers given by listeners (approximately 15%) were replaced by new responses from other listeners. The median questionnaire duration was just over 16 minutes for all listeners. No restrictions were placed in terms of regions in the four countries where the experiment was conducted. Table 2 shows the number of completed questionnaires per listener group, the listener groups’ mean age, sex, self-reported English fluency, and education level.

Table 2 Listeners (N=2266): age, % sex, self-reported English fluency, education level

<table>
<thead>
<tr>
<th>Country</th>
<th>Netherlands (N=567)</th>
<th>Germany (N=617)</th>
<th>Spain (N=540)</th>
<th>Singapore (N=542)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age:</td>
<td>38 (Min = 18; Max = 83)</td>
<td>38 (Min = 19; Max = 83)</td>
<td>37 (Min = 18; Max = 64)</td>
<td>34 (Min = 18; Max = 80)</td>
</tr>
<tr>
<td>Male: Female:</td>
<td>39.7% 60.3%</td>
<td>47.9% 52.1%</td>
<td>36.1% 63.9%</td>
<td>44.6% 55.4%</td>
</tr>
<tr>
<td>Mean self-reported English fluency (SD)</td>
<td>3.61 (SD = .59)</td>
<td>4.03 (SD = .51)</td>
<td>3.62 (SD = .50)</td>
<td>4.21 (SD = .69)</td>
</tr>
<tr>
<td>Education:</td>
<td>12.2% 59.5%</td>
<td>0.0% 42.1% 52.0%</td>
<td>0.0% 55.0% 40.5%</td>
<td>0.2% 84.4% 12.3%</td>
</tr>
<tr>
<td>Undergraduate/Bachelor:</td>
<td>19.5% 2.4% 6.4%</td>
<td>5.8% 0.0%</td>
<td>4.5% 0.0%</td>
<td>2.0% 1.1%</td>
</tr>
<tr>
<td>Master:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Doctorate:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

?Mean 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 a L1 speaker).

2.4 Speaker evaluations and assigned nativeness

Speaker evaluations and the listeners’ estimation of the speakers’ assigned nativeness (Q: ‘Where do you believe the speaker is from?’) were assessed for each speaker. Based on Nejjari et al. (2012, 2021, 2020) and Grondelaers et al. (2015), 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) to what extent they believed the speaker possessed 11 personality traits, representing three dimensions: status (competent, educated, having authority, intelligent and cultured), affect (considerate, pleasant and 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., Zahn & Hopper 1987; Grondelaers et al. 2015). 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 3.

Table 3 Rotated Factor Matrix: factor loadings scores on 11 scales with three factors for listener groups (Netherlands, Germany, Spain, Singapore). Only loadings >.600 are reported.

<table>
<thead>
<tr>
<th></th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Educated</td>
<td>.825</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intelligent</td>
<td>.780</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cultured</td>
<td>.780</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Competent</td>
<td>.723</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confident</td>
<td>.610</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friendly</td>
<td></td>
<td>.843</td>
<td></td>
</tr>
<tr>
<td>Pleasant</td>
<td></td>
<td>.725</td>
<td></td>
</tr>
<tr>
<td>Considerate</td>
<td></td>
<td></td>
<td>.709</td>
</tr>
<tr>
<td>Energetic</td>
<td></td>
<td></td>
<td>.825</td>
</tr>
<tr>
<td>Enthusiastic</td>
<td></td>
<td></td>
<td>.657</td>
</tr>
<tr>
<td>Authoritative</td>
<td></td>
<td></td>
<td>.680</td>
</tr>
</tbody>
</table>

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).

Statistics

Descriptives and frequencies were calculated to establish means and percentages of listener characteristics and responses (Table 2). The factor analyses we applied to trace underlying dimensions in speaker evaluations were principal component analyses, with varimax rotation (SPSS, release 27). We used lme4, R package, 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 describe to what extent the four listener group’s (Dutch, German, Spanish, Singaporean) assigned the matched- and verbal-guise speakers’ native English speaker status (assigned nativeness) (3.1). Secondly, we discuss the impact of perceiving speakers as (non-) native English speakers (assigned nativeness) on speaker evaluations (status, affect, dynamism) (3.2). Lastly, we discuss the listeners’ responses to the individual speakers to understand the effects of voice and accent on speaker evaluations and of assigned nativeness on voice and accent (3.3).

3.1 Assigned nativeness: who is the real native?

Figure 1 shows the assigned nativeness per speaker and per listener group (country) and Table 5 shows the
specific country of origin answers that were given per speaker and listener group (see Supplementary Materials 2 [https://surfdrive.surf.nl/files/index.php/s/KXWnYEICmDyoFIO]). Figure 1, with substantial overlap between most scores for the native English guises, shows that the matched-guise speaker was able to convincingly produce native English accents, according to all listener groups (matched guises: standard British English > 84% within each listener group; standard American English > 86% within each listener group). In addition, the majority of these listeners listener groups also correctly indicated that the standard British and American English matched-guises were British and American English accents. The verbal guise speakers of the native English accents were also convincing as native English speakers according to all listener groups, with the standard British English verbal guises slightly more so than the standard American English verbal guises (standard British English > 82%; standard American English > 69%).

Figure 1 Assigned nativeness %, per speaker, listener group *, 95% confidence intervals

As shown in Plot 1, The Dutch-accented English matched guises were assigned native English speaker status to some degree by all listener groups, most so by Spanish listeners and least so by Dutch listeners (non-native > 94% of listeners). Figure 1 also shows substantial overlap in the assigned nativeness scores for the Dutch-accented matched and verbal guises, meaning that the Dutch listeners viewed all guises similarly. While most German and Singaporean listeners also believed they were listening to non-native English speech, at least 30% of them, and 55% of Spanish listeners, believed they were listening to native English speech. When listeners believed they were listening to a native English speaker, they indicated that the native English speaker was from either the United Kingdom, the United States or Australia. When listeners believed they were listening to non-native 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, the Dutch-accented English verbal guises were recognized best by Dutch listeners, albeit to a lesser degree compared with the matched guises, with over 10% of Dutch listeners believing they were listening to a native 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 a native English speaker. Verbal guise speaker 2 was viewed by at least 30% of German and Singaporean listeners as native English and 55% of Spanish listeners as native English. A pattern similar to that found for our Dutch English matched-guise speaker. In general, Spanish listeners had the most difficulty distinguishing
between native and non-native English accents. The native English assigned verbal guises were viewed as either originating from the United Kingdom, the United States, or Australia. The non-native English assigned verbal guises were mostly assigned Dutch or European origins. Interestingly, some Singaporeans assigned Indian origins to the Dutch-accented matched guises. Similarly, in 10% to 16% of cases, German, Spanish, and Singaporean listeners believed they were listening to Indian-accented English when listening to verbal-guise speaker two (V2 DE).

3.2 The impact of assigned nativeness on speaker evaluations

Mixed model analyses were conducted to see to what extent the speaker evaluations (status, affect, dynamism) were impacted by assigned nativeness, and whether this differed per listener group. The means of the speaker evaluations and the frequencies for assigned nativeness per speaker, accent, and listener group are given in Figures 2-7 and Table 6 (see Supplementary Materials https://surfdrive.surf.nl/files/index.php/s/rlDyr79PW0Z9Uoi).

Status and assigned nativeness

A linear mixed model analysis showed that all effects were significant (Figures 2,3). There was a three-way interaction between speaker, assigned nativeness, and listener group (F(18, 6220.4) = 6.78, p = .00). All three two-way interactions were significant as well: speaker and assigned nativeness (F(6, 6193.9) = 3.00, p = .01), assigned 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), assigned nativeness (F(1, 6399.1) = 59.29, p = .00), and listener group (F(3, 4815.5) = 32.64, p = .00). Figures 2 and 3 also show that the status evaluations for speakers that were considered native or non-native English speakers showed substantial overlap within and between listener groups, suggesting that regardless of listener group, there were generally not many differences between speakers in terms of the status they were assigned on the basis of their assigned nativeness. What is striking is that for all listener groups the speakers that were viewed as non-native English speakers, regardless of their actual English accent, displayed more varied status evaluations per speaker, resulting in wider confidence intervals compared with speakers that were viewed as native 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 a non-native 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 a native English speaker by Dutch listeners, his status evaluations increased significantly, but with a much wider confidence interval. Furthermore, in general, assigned nativeness did affect status evaluations, with speakers that were assigned the position of being a native English speaker (M = 3.87, SD = .63) being attributed higher status than speakers who were not assigned nativeness (M = 3.57, SD = .73) (F(1, 6340.7) = 46.90, p = .00).

Figure 2 Status speakers assigned non-nativeness, per listener group*, 95% confidence intervals
A linear mixed model analysis (Figures 4, 5) showed that there was no three-way interaction between speaker, assigned nativeness, and listener group \((F(18, 6351.4) =1.39, p =.13,\) ), as well as no two-way interaction assigned nativeness and listener group \((F(3, 6351.0) =1.29, p =.28,\) ). However, there was a two-way interaction between speaker and assigned 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 within speakers when they were assigned native English status, or not, and affect evaluations for individual speakers also
differed between listener groups, even though listener group in general did not impact affect evaluations (F(3, 5027.5) =1.31, p =0.27). The affect evaluations were, however, impacted by assigned nativeness (F(1, 6396.3) =4.11, p =.04). Assigning a native English speaker (both matched guises and verbal guises) the status of native English speaker led to more consistent affect evaluations, observed in narrower confidence intervals compared to speakers assigned non-native English speaker status (Figures 4, 5), and affect was higher for speakers viewed as native English (M =3.53, SD =.76) compared to speakers who were not (M =3.38, SD =.74).

![Figure 4 Affect speakers assigned nativeness, per listener group*, 95% confidence intervals](image)

*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.

![Figure 5 Affect speakers assigned non-nativeness, per listener group*, 95% confidence intervals](image)

*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.
**Dynamism and assigned nativeness**

A linear mixed model analysis (Figures 6, 7) showed that there was no three-way interaction between *speaker*, *assigned nativeness*, and *listener group* ($F(18, 6229.1) = 1.47, p = .09$), and no two-way interaction between *speaker* and *assigned nativeness* ($F(6, 6221.3) = 1.51, p = .17$). There was a two-way interaction between *assigned nativeness* and *listener group* ($F(3, 6332.3) = 3.09, p = .03$), between *speaker* and *listener group* ($F(18, 6125.4) = 1.74, p = .03$). This means that *dynamism* evaluations varied significantly between *listener groups* in general, but also for an individual *speaker* when they were *assigned nativeness*. 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 native English speaker status. This might mean that the correct combination of accent, *assigned nativeness*, and a native (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, *assigned nativeness* impacted *dynamism* ($F(1, 6404.9) = 11.40, p = 001$) in that speakers that were assigned native English speaker status ($M = 3.36, SD = .79$) were viewed as more dynamic than speakers who were not viewed as native English speakers ($M = 3.14, SD = .81$). A native English accent produced by both a matched-guise and verbal-guise speaker that was assigned native English speaker status was given more consistent *dynamism* evaluations compared to that same native English speaker when they were considered to be a non-native English speaker. Figure 7 also shows more consistent *dynamism* evaluations for the Dutch English speakers that were viewed as non-native English speakers compared to the native English speakers that were viewed as non-native English speakers.

**Figure 6 Dynamism speakers assigned nativeness, per listener group*, 95% confidence intervals**

![Figure 6 Dynamism speakers assigned nativeness, per listener group*, 95% confidence intervals](image)

*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.

**Figure 7 Dynamism speakers assigned non-nativeness, per listener group*, 95% confidence intervals**

![Figure 7 Dynamism speakers assigned non-nativeness, per listener group*, 95% confidence intervals](image)
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.

In summary, our analyses clearly show that speaker evaluations are significantly affected by assigned nativeness, for all dimensions: status, affect, dynamism. What is interesting is that the status, affect, and dynamism evaluations of (non-)native English speakers, regardless of assigned nativeness, mostly did not differ between speakers and listener groups. However, for the Dutch listener group, when the Dutch-accented English speakers were assigned the status of non-native English speaker this led to significantly lower status when they were identified as native English speakers of Dutch. Affect and dynamism were equally significantly affected by assigned nativeness since (correctly) assigning nativeness led to the most consistent affect and dynamism evaluations. This consistency implies that when the listener groups, who each 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 further investigate the impact of voice and accent on speaker evaluations.

3.3 The impact of voice and accent on speaker evaluations

To further understand to what extent voice impacts responses to speech samples we compared the speaker evaluations to the matched guises with the speaker evaluations to the verbal guises. All statistical analyses were done with linear mixed models, as reported in the previous sections (Supplementary Materials 3https://surfdrive.surf.nl/files/index.php/s/r1Dyrg79PWoZ9Uoi for means and SDs).

Table 7 shows the effects of accent and assigned nativeness perspeaker (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 a native 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, assigned nativeness also positively influenced status, meaning that when Dutch listeners assigned the matched-guise speaker the status of being a native English speaker, regardless of the actual accent he produced, this positively influenced status evaluations of the matched-guise speaker. Furthermore, for the matched-guise speaker, accent positively influenced affect, meaning that the matched-guise speaker aroused affect evaluations on the basis of the produced (non-)native English accents with
significantly higher affect evaluations for the standard British English and Dutch-accented English matched guises, compared with the standard American English matched guises. In terms of dynamism, a positive effect of assigned nativeness was observed for the verbal guises. There was no combined effect of accent and assigned nativeness on speaker evaluations for either verbal or matched guises.

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

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 DE) being perceived as having a significantly higher status than the other (V1 DE). 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. Similar to 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 assigned nativeness and voice for status, which meant that when a speaker was considered a native 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 BE for standard British English). Finally, Singaporean listeners

Table 7 Effects (positive = +, negative = -, or no effect = 0) of accent, assigned nativeness 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.

<table>
<thead>
<tr>
<th>Listener group</th>
<th>Factor</th>
<th>Status</th>
<th>Affect</th>
<th>Dynamism</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MG VG</td>
<td>MG VG</td>
<td>MG VG</td>
<td>MG VG</td>
</tr>
<tr>
<td>NL</td>
<td>Accent</td>
<td>+ +</td>
<td>+ 0</td>
<td>0 0</td>
</tr>
<tr>
<td></td>
<td>Assigned nativeness</td>
<td>+ 0</td>
<td>0 0</td>
<td>0 0</td>
</tr>
<tr>
<td></td>
<td>Accent x Assigned Nativeness</td>
<td>0 0</td>
<td>0 0</td>
<td>0 0</td>
</tr>
<tr>
<td>Germany</td>
<td>Accent</td>
<td>0 0</td>
<td>0 0</td>
<td>0 0</td>
</tr>
<tr>
<td></td>
<td>Assigned nativeness</td>
<td>0 0</td>
<td>0 0</td>
<td>0 0</td>
</tr>
<tr>
<td></td>
<td>Accent x Assigned Nativeness</td>
<td>0 0</td>
<td>0 0</td>
<td>0 0</td>
</tr>
<tr>
<td>Spain</td>
<td>Accent</td>
<td>0 0</td>
<td>0 0</td>
<td>0 0</td>
</tr>
<tr>
<td></td>
<td>Assigned nativeness</td>
<td>0 0</td>
<td>0 0</td>
<td>0 0</td>
</tr>
<tr>
<td></td>
<td>Accent x Assigned Nativeness</td>
<td>0 0</td>
<td>0 0</td>
<td>0 0</td>
</tr>
<tr>
<td>Singapore</td>
<td>Accent</td>
<td>0 0</td>
<td>0 0</td>
<td>0 0</td>
</tr>
<tr>
<td></td>
<td>Assigned nativeness</td>
<td>0 0</td>
<td>0 0</td>
<td>0 0</td>
</tr>
<tr>
<td></td>
<td>Accent x Assigned Nativeness</td>
<td>0 0</td>
<td>0 0</td>
<td>0 0</td>
</tr>
</tbody>
</table>

Accent and assigned nativeness appeared to have had relatively limited effects on speaker evaluations for especially 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, assigned 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 DE) being perceived as having a significantly higher status than the other (V1 DE). 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. Similar to 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 assigned nativeness and voice for status, which meant that when a speaker was considered a native 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 BE for standard British English). Finally, Singaporean listeners
only showed a preference for a speaker for affect, which was positively impacted when an individual speaker was viewed as both a native English speaker (assigned nativeness) and had a voice quality that most likely appealed to the listener.

5. Conclusion and discussion

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

RQ1 Assigning nativeness in an ELF context and its impact on speaker evaluations

Firstly, we wanted to determine to what extent non-native English listener groups assign standard British and American English accents and Dutch-accented English (non-)nativeness, and whether assigning (non-)nativeness to speakers impacts speaker evaluations.

What our results indicate is that in general all four listener groups (Dutch, German, Spanish, Singaporean) were good at assigning the two standard English accents, and especially the standard British English accent, native English speaker status, even though they did not always know which individual native English accent they had listened to. These results suggest that, similar to native (English) listeners, both native and non-native speakers of English can distinguish well between (non-)native English accent varieties, but not necessarily between native English accent varieties (Nolan 2003; Brunner 2009; Gnevsheva 2018; McGorrery & McMahon 2017; Wong and Babel 2017). This may be because while native English accents are significantly similar phonetically, they are phonetically significantly different from non-native English accents (see also McMahon, Heggarty, McMahon & Maguire 2007). These results also suggest that our non-native English listeners were likely familiar with at least one of the native English accent varieties. Indeed, the best recognized accent variety in our study was British English, which is the English variety traditionally used in general education and commercial language training in Europe (The Netherlands, Germany, Spain), and the English variety historically connected to Singapore and still being used in Singaporean teacher training (British Council, 2022).

In terms of identifying the L2 English accent, it is striking that when the Dutch-accented English speech samples were considered non-native English, they were viewed by all listener groups as having European origins, even though the content of the speech samples in no way alludes to a European setting or includes 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 country of origin of the speaker.

<table>
<thead>
<tr>
<th>Verbal Guises</th>
<th>Factor</th>
<th>Status</th>
<th>Affect</th>
<th>Dynamism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Netherlands</td>
<td>Predictor</td>
<td>Assigned nativeness</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Assigned nativeness x Assigned nativeness</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Germany</td>
<td>Predictor</td>
<td>Assigned nativeness</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Assigned nativeness x Assigned nativeness</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Spain</td>
<td>Predictor</td>
<td>Assigned nativeness</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Assigned nativeness x Assigned nativeness</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Singapore</td>
<td>Predictor</td>
<td>Assigned nativeness</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Assigned nativeness x Assigned nativeness</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Interestingly, between 9 and 16.2% of German, Spanish and Singaporean listeners believed that the accent origin of one Dutch-accented English verbal-guise speaker (V2 DE) was Indian. More generally, some Singaporeans indicated that the Dutch-accented English accent matched and verbal guises (5.4%, 5.6%, 12.4%) were originally Indian. This might be because the Singaporean listeners who believed that the speech samples were non-native English accents compared them to the non-native English accents they are themselves familiar with. Singapore is a multi-ethnic society whose residents are mostly of Chinese descent, followed by Singaporeans that are of Malay or Indian descent (less than 10%) on the basis of paternal data (Department of Statistics Singapore 2022). The Singaporean listeners might have perceived the Dutch-accented English accent as phonetically and prosodically most similar to an Indian-accented English, which they could be relatively familiar with 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 native English by three of the non-native English listener groups (German, Singaporean, Spanish), who were, with the exception of German listeners, most likely not 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 native English speakers, in most cases British English speakers. In contrast, the Dutch listeners displayed their familiarity with Dutch-accented English since 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 matched and verbal guises.

In terms of assigning (non-)nativeness to the speakers and its impact on speaker evaluations, we conclude that (correctly) perceiving someone as a native English speaker leads to more consistent speaker evaluations (status, affect, dynamism) compared to perceiving someone (correctly) as a non-native English speaker, even if the specific native English accent is not (correctly) identified. In addition, being viewed as a native English speaker seems to positively influence speaker evaluations on 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 German, Spanish, and Singaporean listeners, with Dutch listeners assigning significantly lower status when they correctly identified Dutch-accented English as non-native English and as coming from a native speaker of Dutch. It appears that for non-native English listeners, like for native English listeners (see Nejjari et al. 2012), sufficient familiarity with a specific non-native 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 has low social status compared with native English, because they have very specific associations, based on experience and exposure, with Dutch-accented English and its speakers. It is conceivable therefore, that a similarly strong response to Dutch listeners’ “own” English accent variety might not be observed for other non-native English listeners and “their” English accent varieties. Unlike status, affect and dynamism, evaluations were more consistent for Dutch-accented English speakers that were identified as non-native English speakers than for Dutch-accented English speakers that were assigned native English speaker status. This was the case for all four listener groups.

What the results in general suggest is that an essential part of speaker evaluations is the estimation of a speaker’s nativeness, and that, when this estimation is correct, speaker evaluations are most consistent. Also, 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 non-native English listeners is a norm that associates nativeness (versus 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 (Nejjar et al. 2021; Canagarajah 2007). It is not so much that accentedness does not matter to non-native English speakers in ELF contexts, but it appears that non-native English speakers mistakenly categorize English accents, and this can make it seem as if their responses reflect a non-traditional language norm, when it does not.
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, so we assessed whether listeners responded differently to matched guises compared to verbal guises in terms of accent, assigned nativeness and speaker evaluations (status, affect, dynamism). This also allowed us to compare the effectiveness of matched guises and verbal guises in accentedness research focused on (non-)native English accents and to investigate to what extent 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 assigned 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 DE) being perceived as having a significantly higher status than Dutch English verbal-guise speaker 2 (V2 DE). 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. So, despite knowing the language background of V1 DE, the Dutch listeners assigned him significantly higher status compared to V2 DE, which shows how most likely voice impacted his status evaluations. In addition, affect and dynamism were also positively affected by voice, with almost all verbal-guise speakers within one accent 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 observe the impact of a speaker’s voice again in the preference of most listener groups for one verbal-guise speaker of one accent over the other verbal guise speaker of that same accent, for the tested native English accents. Furthermore, when it comes to status, this study shows that it seems to pay to be viewed as a native English speaker and having a voice with qualities that lead to positive evaluations, as the results for German listeners showed us. Future research could investigate whether similar patterns would emerge for a wide range of (non-)native English listener groups, for example on the basis of Kachru’s concentric circles model, and which voice(s) qualities evoke positive speaker evaluations in which 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 speaker evaluation processes are based on estimations of a speaker’s language background and that perceived nativeness functions as 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 a person’s voice quality happens to stand out in a positive or negative way to listeners, the subsequent outcome may be that a speaker is evaluated more positively or negatively, despite their (non-)nativeness. This also suggests that voice, like accent, is a fundamental aspect of speaker evaluations.

Limitations, implications, and future research

In our study, we tested two native English varieties and one non-native English variety from a language that is part of the same language family as English, namely the West-Germanic language family. Although our listeners were not able to identify the origin of the Dutch-accented English speech samples, they mostly did assign them European origins and in quite a number of cases even native 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, nor assessed how confident listeners felt about their answers on the country of origin of the speakers. We assumed that the German, but especially Spanish and Singaporean listeners, were not familiar with Dutch and Dutch-accented English, but we did not ask this. Even for the German listeners, who were the best at identifying the Dutch-accented English (40%), we did not assume their familiarity since the experiment was conducted in all of Germany, not necessarily in the border area between the Netherlands and Germany.

More knowledge on assigning (non-)nativeness can give us a better understanding of the role familiarity
plays and when an accent variety is considered native or not, for instance, on the basis of specific phonetic features. Subsequently, we can better understand whether accent varieties from certain language families are also recognized as coming 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 non-native English accents, their phonetic and prosodic features, and how these are categorized and evaluated by a variety of non-native English listeners.

In terms of employing the matched-guise and verbal-guise technique in our study, we can confidently state that matched guises offer a valid method to investigate how listeners assess 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 so 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, compared to the other British English verbal guise speaker, but also compared to all other speakers. In terms of the application of the matched-guise technique to assess responses to both native and non-native English accents, it will remain challenging to find matched-guise speakers who can representatively produce a combination of (non-)native English accents, let alone from several language families (as discussed in Nejjar 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 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 cultural or even linguistic background (e.g., Majid 2012). Future research could zoom in on this further 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 non-native 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. Foreign language education therefore should focus on developing more tolerance to varieties of non-native Englishes to ensure unbiased evaluations of the abilities and character of non-native English speaker groups.

References


