frontiers | Frontiers in Language Sciences

TYPE Original Research
PUBLISHED 08 January 2024
DOI 10.3389/flang.2023.1292344

OPEN ACCESS

EDITED BY
Susanne Brouwer,
Radboud University, Netherlands

REVIEWED BY
Anna Lorenzoni,
University of Padua, Italy
Václav Jonáš Podlipský,
Palacky University Olomouc, Czechia

*CORRESPONDENCE
Thanh Lan Truong
✉ thanh-lan.truong@uni-tuebingen.de

RECEIVED 11 September 2023
ACCEPTED 11 December 2023
PUBLISHED 08 January 2024

CITATION
Truong TL and Weber A (2024) Investigating the effect of nativeness and speaker age on the credibility of spoken sentences.
Front. Lang. Sci. 2:1292344.
doi: 10.3389/flang.2023.1292344

Investigating the effect of nativeness and speaker age on the credibility of spoken sentences

Thanh Lan Truong* and Andrea Weber

Psycholinguistics and Applied Language Studies, English Department, University of Tübingen, Tübingen, Germany

Previous studies found mixed results regarding a bias in credibility ratings for trivia statements made by L2 speakers in comparison to L1 speakers. Perceptual fluency, social attitudes, and pragmatic lenience have been proposed as underlying causes for the bias. The present study examined credibility ratings for L2 speakers and extended the scope of the investigation by adding the factor age of the speaker. In the present study, German native adult listeners were asked to judge the credibility of trivia statements recorded by L2 adults and L1 children in Experiment 1 and statements by L1 adults and L1 children in Experiment 2. In Experiment 1, the ratings showed no difference in credibility between the L2 adult statements and the L1 child statements. In Experiment 2, listeners rated statements from L1 child speakers as more credible than statements from L1 adult speakers, suggesting a positive bias for the credibility of children. The results are discussed in terms of their relevance for previously suggested causes contributing to a credibility bias.

KEYWORDS

credibility, non-native speech, child speech, perceptual fluency, attitude, pragmatic lenience

1 Introduction

In our daily lives, we communicate with a diverse range of interlocutors, including family members and strangers, speakers of different gender, young children and elderly speakers, and speakers with the same or a different language background. Regardless of who our interlocutors are, as listeners we not only have to figure out the linguistic meaning of what has been said, but we also often have to evaluate the validity, accuracy, and overall merit of what has been said for conversation. Imagine a young son engaging in a conversation with his father, who asserts that the K2 is the highest mountain in the world. The son might accept this statement as credible and accurate, and in an ensuing conversation they might discuss the Karakoram range instead of the Himalayan range, where Mount Everest, the true highest peak, is situated. On the other hand, if the young son or a complete stranger with a strong non-native accent makes the same statement, the father might judge the veracity of the statement differently.

Numerous studies have investigated credibility from different perspectives with a focus on comparing ratings for utterances made by non-native (L2) speakers with that for native (L1) speakers. Encountering individuals who are not native speakers of the local native language has become increasingly normal in our multicultural and globalized world. Whether it is in a social setting, educational institution, or workplace, listening to foreign-accented speech has become an integral part of our everyday interactions, and studying how foreign-accented utterances are judged in terms of credibility has become a worthwhile endeavor. In a seminal study, Lev-Ari and Keysar (2010) have found that native adult listeners judge English trivia statements such as "*ants don't sleep*" as less credible when they

have been produced by L2 speakers with a non-native accent in comparison to when they have been produced by native speakers. This effect was found for both mildly and heavily accented L2 speakers with various language backgrounds. When the native adult listeners were informed beforehand about this effect, they no longer rated mildly accented statements as less credible, but strongly accented statements were still judged to be less credible than native statements. Foreign-accented speech typically deviates from the norms of a target language, and these deviations can make it more difficult for native listeners to process the speech signal. Such a reduction in perceptual fluency for L2 speech in comparison to L1 speech has been found repeatedly in the literature (for a review see, Cristia et al., 2012). For example, Bürki-Cohen et al. (2001) showed longer reaction times when words were produced by non-native speakers compared to when they were produced by a native speaker, which demonstrates an influence on reduced perceptual fluency. At the same time, L2 speakers are often met with negative attitudes and stereotypes (Dixon et al., 2002). Lev-Ari and Keysar (2010) discussed both aspects as being potential causes for a bias against L2 speakers' credibility, but interpreted their findings as favoring an explanation based on perceptual processing difficulties, also because, in an attempt to exclude negative attitudes, they had told participants beforehand that the statements were only recited by the speakers and, in fact, been written by the native experimenter.

The study by Lev-Ari and Keysar (2010) spurred further investigations that examined credibility judgments made by native listeners from various angles, including, for example, the credibility of speakers with regional accents (Frances et al., 2018) and diverse language contexts (e.g., De Meo, 2012; Podlipský et al., 2016; Hanzlíková and Skarnitzl, 2017; Stocker, 2017). It is important to note that these studies often differed in their methodology. They employed, for instance, distinct experimental approaches by tackling the issue from a neurolinguistic perspective with ERP measures (Foucart et al., 2019, 2020; Foucart and Hartsuiker, 2021) or embedded the stimuli from native speakers in noise at different signal-to-noise ratios (e.g., Souza and Markman, 2013). The hypothesis for the noise manipulation was that, if processing difficulties lead to a negative bias, then noise should have a similar effect as the accent of L2 speakers, as noise is known to increase processing difficulties, too. Noise did, however, not affect the credibility ratings for native speakers in Souza and Markman (2013). Podlipský et al. (2016) embedded stimuli from native and non-native speakers in noise and found no bias in credibility ratings of native listeners. Only reaction times of the ratings correlated mildly with the ratings themselves which was interpreted as limited support for the hypothesis that decreased perceptual fluency leads to decreased credibility. It must be noted, however, that Podlipský et al. (2016) distributed statements and speakers across conditions, such that statements were attributed to specific speakers. In contrast, Lev-Ari and Keysar (2010) designed a within-participant and within-item design, meaning that both speakers and statements were fully crossed across conditions, which ensured that the credibility of individual statements could not have influenced the observed pattern of result.

While recent results by Boduch-Grabka and Lev-Ari (2021) supported the finding by Lev-Ari and Keysar (2010) that listeners trust information less when it is delivered with an L2 accent, Souza and Markman (2013) found, for example, no negative bias for the credibility of utterances by L2 speakers. Others have reported partial effects (see also Podlipský et al., 2016; Hanzlíková and Skarnitzl, 2017; Stocker, 2017) and some even found that statements with strong accents were considered to be more credible rather than less credible (Fairchild and Papafragou, 2018; Frances et al., 2018; Lorenzoni et al., 2022). Frances et al. (2018) compared credibility ratings for various regional accents (i.e., non-standard accents) and local accents (i.e., standard accents) and found that statements with a stronger regional accent were considered to be more credible. Such a positive bias for non-standard accents can be considered to be in line with findings from Fairchild and Papafragou (2018). Fairchild and Papafragou (2018) asked native participants to rate how much sense written under-informative sentences such as "*some people have noses with two nostrils*" make. The sentences were presented in written form, and speaker bios for either a native or non-native speaker accompanied the sentences. It was found that these under-informative written sentences were judged to make more sense when the sentences could be attributed to non-native speakers, as opposed to when the speaker was supposedly native. Lorenzoni et al. (2022) confirmed such a positive bias for credibility ratings (Fairchild and Papafragou, 2018), showing that native participants were more inclined to judge written trivia statements as true when they could be attributed to non-native speakers rather than to native speakers. Processing difficulties could not directly account for the findings, since stimuli were presented in identical written form, and the authors rather argued for an effect of pragmatic lenience toward non-native speakers that is related to participants' beliefs about these speakers' lower linguistic competence.

In sum, a bias in credibility ratings toward or against non-native speakers is not always found and different explanations for such a bias have been provided in the literature, in particular processing fluency, attitudes, and pragmatic lenience. Regarding processing fluency, it is well-attested that the ease or difficulty with which information is processed can influence judgments (e.g., Schwarz, 2004; Oppenheimer, 2008). For example, statements that feel easy to process are perceived as more pleasant (Reber et al., 2004). L2 speech deviates from standard native norms both in terms of segmental and suprasegmental deviations but also in the amount of variability in comparison to native speech (e.g., Wade et al., 2007). While listeners can quickly (Flege, 1984) and reliably identify non-nativeness in speakers (Munro et al., 2010), listening to L2 speech does come with its challenges. It has been found again and again to be more challenging than listening to native speech (for a review see, Cristia et al., 2012) and this reduced perceptual fluency could affect listeners' credibility evaluations of L2 speech (Dragojevic and Giles, 2016; Dragojevic et al., 2017). Lev-Ari and Keysar (2010) argued indeed in their seminal study for an account of reduced processing fluency for the observed negative credibility bias. Later, in Boduch-Grabka and Lev-Ari (2021), the authors promoted a link between attitudes and processing fluency, since even brief exposure to a foreign accent in their study reduced the negative bias in credibility ratings due to an improvement in the processing of the accent.

Stereotypes about L2 speakers, which typically promote prejudice and negative attitudes, could also have an impact on the credibility of utterances by L2 speakers. That is, non-native speakers are often met with negative attitudes which possibly are being promoted by in-group biases and not necessarily by the accent as such. Accent could simply serve as marker for the biases (e.g., Dixon et al., 2002) and result in spontaneous and implicit social categorization, helping to identify an individual as either an in-group or an out-group member (Pietraszewski and Schwartz, 2013). As a consequence, non-native speakers regularly have to face stigmatization, social ostracism, and unfair jurisdiction (e.g., Dixon et al., 2002), and are often rated lower than native speakers in terms of intelligence, attractiveness, education, kindness, and prestige (Edwards, 1999; Fuertes et al., 2002; Lindemann, 2003; Munro et al., 2006; Anderson et al., 2007; Gluszek and Dovidio, 2010). These social attitudes could negatively impact judgments of their credibility, too. On the other hand, non-native speakers are also given the benefit of the doubt when their utterances are being interpreted (e.g., Gibson et al., 2017). Indeed Fairchild and Papafragou (2018) and Lorenzoni et al. (2022) found a positive bias toward utterances of non-native speakers and argued that pragmatic lenience toward L2 speakers accounted for this finding. Following this account, non-native speakers more frequently receive forgiveness for their lack of linguistic competence, and pragmatic lenience makes native participants accept more readily sentences that are under-informative or the credibility of trivia statements made by non-native speakers. Corroborating evidence for this account comes from electrophysiological studies (for discussion see, Foucart et al., 2020).

The diverging pattern of results across studies makes it impossible for a single factor, such as processing fluency or pragmatic lenience, to fully account for the credibility judgments for non-native speakers. It is more likely that several factors are at play, possibly with varying strength of influence in various contexts. In the present study, we took a new approach by exploring credibility ratings not only for L2 adult speech but also for L1 child speech. To our knowledge, no one has tested yet the credibility of trivia statements made by children. The closest related evidence comes from courtroom studies. These studies have shown, for example, that judges perceive children as more honest than adult witnesses, despite their limited memory capacities and verbal skills, which can make them appear less reliable than adults in a courtroom situation (Ross et al., 1990; Bala et al., 2005). This is starkly different from the situation for L2 adult speakers. For example, L2 speakers are usually attributed more guilt than native defendants in a mock trial study conducted by Seggie (1983) and Romero-Rivas et al. (2021). Most importantly, not only are those individuals who are accused of crimes impacted, but also those who provide evidence for testimony. Testimony by L2 speakers can face varying degrees of scrutiny regarding the validity and reliability of their statements. Frumkin and Stone (2020), for instance, found that eyewitnesses with L2 accents were perceived as less credible than those with L1 accents.

While both L2 adult speakers and L1 children typically deviate in their speech from the adult norms of the target language, children are typically met with positive attitudes (Ross et al., 1990; Bala et al., 2005). Indeed, social studies have examined age differences in implicit and explicit attitudes and found overall more positive associations for younger people (Axt et al., 2014; Chopik and Giasson, 2017). For example, Axt et al. (2014) found that social evaluations follow a specific hierarchy, with the strongest positive associations observed for children, followed by a steady decrease with age (i.e., children > young adults > middle age adults > older adults). It is conceivable that these positive associations for children positively impact credibility ratings (Ross et al., 1990; Bala et al., 2005).

In spoken language, the voice is the major medium that conveys important information about a speaker such as gender, age, height and weight, and ethnic background (Latinus and Belin, 2011). As children generally have higher-pitched voices, it is easy for listeners to differentiate child voices from adult voices. The difference in pitch mainly arises from differences in the length of vocal folds for children and adults. The vocal folds of an 8-year-old child, for example, have grown to approximately 8mm in length, compared to adults, whose average vocal fold length ranges from 12 to 21mm. The fundamental frequency (i.e., F0) of an infant's voice is at birth around 500 Hz high. As the larynx grows with age, F0 drops to about 275 Hz by the age of eight, while the F0 range remains quite stable throughout childhood, at about 2.5 octaves, F0 decreases progressively as children mature (Vorperian and Kent, 2007; but see also Kreiman and Sidtis, 2011). Anatomical differences in the vocal-tract geometry also change the spectral characteristics of child speech (at least up to age 13) in comparison to adult speech (e.g., Kent, 1976; Potamianos and Narayanan, 2003), and the developmental aspect of language acquisition leads to greater variability in the signal with more deviations and variability being present the younger the children are (Tingley and Allen, 1975; Smith et al., 1983). In this sense, acoustic and linguistic properties of L1 children's speech are distinct from those of L1 adult speech (Lee et al., 1999). Based on these differences, listening to child speech could be harder than listening to native adult speech. A recent study on speaker identification by Cooper et al. (2020) indeed found that native adult listeners need more time and are less accurate at learning to identify child speakers than adult speakers. Since both L1 children and L2 adult speakers can be regarded as having lower linguistic competence, native listeners might be pragmatically lenient toward both speaker groups because of the deviations in speech from adult target language norms.

In the present study, we compared credibility ratings for trivia statements recorded by L2 adults with a noticeable accent and L1 children, and for trivia statements recorded by L1 adults and L1 children. An account based on pragmatic lenience (Fairchild and Papafragou, 2018; Lorenzoni et al., 2022), would expect credibility ratings of L1 adult listeners to be more lenient, i.e. positive, toward statements made by L2 adult and L1 children in comparison to statements made by L1 adults. Children are, however, typically associated with more positive attitudes than adults. Particularly, L2 adults often must endure stigmatization and are evaluated negatively (Gluszek and Dovidio, 2010), while children regularly encounter caring and kindly reaction (Ross et al., 1990; Bala et al., 2005). If social attitudes significantly influence credibility ratings in the present study, then credibility ratings for utterances recorded by L1 children might be higher than those for both groups of adults. In terms of an influence of perceptual fluency,

the predictions are maybe less clear since perceptual fluency for child speech has not been previously investigated. However, given that both L2 adult speech, for which reduced processing fluency in comparison to L1 adult speech has been amply attested (for an overview see, Cristia et al., 2012 and see also Bürki-Cohen et al., 2001), and L1 child speech is more variable than L1 adult speech, perceptual fluency might be reduced for L1 child speech in comparison, too. We presented German trivia statements like "*Das Auge des Vogel Strauss ist grösser als sein Gehirn*," "an ostrich's eye is bigger than its brain," to German adult participants who had to rate their credibility. The statements were recorded by L1 adult speakers of German, L2 adult speakers of German with various L1 language backgrounds, and L1 child speakers of German who were between 6 and 11 years old. Care was taken that the linguistic competence of both the L2 adult speakers and the L1 child speakers was high enough for them to be able to produce the sentences fluently. We wanted to avoid credibility ratings that are based on misunderstanding of the statements or unclarity about what has been said, and therefore aimed for high intelligibility for all recordings. Note that high intelligibility for all three speaker groups in the sense of listeners being able to correctly understand what the speakers had said (e.g., Munro et al., 2006) does not imply that the three groups were comparable in terms of the ease with which their recordings could be processed, since high intelligibility does not necessitate high perceptual fluency.

2 Experiment 1

2.1 Method

2.1.1 Material

Forty-five trivia statements were taken from Lev-Ari and Keysar (2010) and translated from English into German. The majority of the trivia statements were about the animal kingdom (e.g., "*Das Auge des Vogel Strauss ist grösser als sein Gehirn*," "an ostrich's eye is bigger than its brain." Nine of the statements were replaced with new trivia statements because the German translation did not work well. For instance, "the original name for butterfly was flutterby" was replaced since the German word for butterfly (i.e., *Schmetterling*), does not entail the embedded words "flutter" and "by." The trivia statements were ones for which it was assumed that the correct answer is usually not known for the participants (while in fact, half of them were true and half were false), thereby reducing the likelihood of fixed judgments on the endpoints of a credibility scale. In order to have some statements for which participants could be more certain regarding their truth value, 15 filler statements with obvious truth values were added (e.g., "*Brokkoli ist ungesund*," "broccoli is unhealthy").

The statements were recorded by four L2 adult speakers (mean age = 41, age range = 38–47) and by four L1 child speakers (mean age = 9, age range = 6–11). The correctness of the statements was not shared with the speakers, so that their speech should not have been affected by it. The recording sessions took place in a sound-attenuated room with a high-quality microphone and a sampling rate of 44 kHz. All speakers were recorded separately. The native language background of the L2 adult speakers comprised of Arabic, Chinese, Spanish, and Russian, and all L2 adult speakers had a noticable accent. The L1 child speakers were recorded together with their mothers. The mother first read aloud from orthographic transcription, and the child was prompted to repeat after her.[1] The current study opted to record multiple female speakers, because exploratory results from Truong and Weber (2020), who used a single male adult speaker and a single female child speaker, suggested that vocal qualities of individual speakers possibly impact credibility ratings (for social impressions based on voice see also e.g., Baus et al., 2019). Care was taken, that all recordings were fluent and free of hesitations. The style of speaking was natural with no noticeable hyper- or hypoarticulation. The mean fundamental frequency of both speaker groups for the complete statements was measured with Praat and checked manually (Boersma and Weenink, 2023). The L2 adult speakers had an average F0 of 233 Hz; L1 child speakers had an average F0 of 256.8 Hz. The F0 values between the two groups of speakers (i.e., adults vs. children) were significantly different ($t = 9.3$, $df = 358$, Cohen's $d = 0.9$, $p < .001$). Statements made by the L2 adult speakers were longer on average than statements made by the children (L1 children: 3,389.6 ms, L2 adults: 4,360.2 ms; $t = -8.6$, $df = 358$, Cohen's $d = 0.9$, $p < .001$).

Eight lists with the 45 experimental statements and the 15 filler statements were created. In each list, half of the statements had been produced by L2 adult speakers (a total of 30 statements distributed across four speakers) and half by L1 child speakers (with a total of 30 statements distributed across 4 speakers); individual speakers were counterbalanced across lists using Latin square. The order of trials was randomized once and the same for each participant; participants started the experiment always with two practice statements.

2.1.2 Participants

Forty-nine native listeners of German, between 18 and 49 years old (mean age = 24; SD = 5.9), participated in Experiment 1 for a small monetary reimbursement. All participants grew up monolingually with German, were female, and reported no visual or hearing impairments.

2.1.3 Procedure

The experiment was conducted online and run on Gorilla, which is an experiment builder software and host of online research studies (www.gorilla.sc) (Anwyl-Irvine et al., 2020).

Participants gave electronically informed consent, before they were given detailed written instructions, asking them to sit in a quiet room, to reduce distractions, and to wear over-the-ear or in-ear headphones during the entire experiment. Furthermore, participation was only permitted using a laptop or computer with the browsers Firefox, Google Chrome, or Safari, since the correct functioning of the experiment with these browsers and hardware devices was thoroughly tested in advance by our research assistants. Participants were automatically disqualified if the technical requirements were not fulfilled.

1 Elementary school starts at the age of six in Germany, and by the age of seven reading aloud unprompted is typically still less fluent than for adults. Repeating after an auditory prompt ensured that the children produced the sentences fluently.

TABLE 1 LMER result summary for Experiment 1.

Fixed effects	*b*	SE	*t*	Pr>|*t*|	
(Intercept)	6.11	0.33	17.9	<2e-16	***
Speaker	−0.03	0.14	−0.25	0.8	
Random effects	**Variance**	**SD**			
Participants	0.85	0.92			
Items	4.15	2.03			

***$p<.001$.

To ensure that participants took part in the experiment wearing headphones, a headphone-screening task (i.e., 3AFC paradigm) was incorporated prior to the actual experiment (Milne et al., 2021). Participants heard three intervals of randomly ordered white noise with equal frequency and duration, but one interval contained a Huggins Pitch tone. Participants were asked to identify which of the three tones contained the hidden pitch. The test consisted of six trials and participants needed five correct responses to pass the headphone test in order to proceed to the actual experiment.

All experimental material was loaded prior to the start of the experimental session to ensure no loading delays during the experiment. The experiment started with two practice trials, followed by the 60 trivia statements. Each statement was presented once. Furthermore, participants were asked to genuinely assess the veracity of each statement using a sliding scale. The scale was similar to the one used by Lev-Ari and Keysar (2010): the scale was set to be 14 cm long in Gorilla and the left end of the scale (at point 0) was labeled definitely false' and the right end (at point 14) was labeled "definitely true." The scale was always positioned at the center of the screen, but as screen size and resolution could vary for at-home participants, the true length of the shown scale in terms of cm could diverge to some degree. Participants assessed the veracity level of each statement by dragging the slider bar, initially positioned at the center of the scale for each trial, and moving it to the desired answer position. Although not visible to participants, the positions on the scale were measured with mm precision (yielding 140 possible ratings points), with higher numbers indicating higher perceived credibility. Participants were encouraged to use the full scale. They had to click on the play button to listen to a statement, before they could enter their judgment. In addition, participants were asked to indicate after each judgment, whether they had heard the statement before and already knew the correct answer. Three possible answer options were given: yes, no, and unsure. In addition, participants also had to indicate if they had understood the statement (binary response options: yes or no).

2.2 Results

R Core Team (2021) (version 4.0.5) and lme4 (Bates et al., 2015) were used to perform a linear mixed effects analysis on listeners' perceived credibility. Since we expected that known statements are less likely to be affected by speaker attributes, only statements that were unknown to the participants were included in the analysis (80.4% of the data).[2] Participants had indicated for every statement that they had understood it and no further data points were removed from the analysis.

The model included *ratings* as the dependent variable and *speaker* (L2 adult and L1 child, sum coded as 0.5 and −0.5, respectively) as an independent variable. The random effect structure included *by-item* varying intercepts, and *by-participant* varying intercepts and varying slopes for *speaker*.[3] The analysis showed no significant effect of *speaker* ($b = -0.03$, $SE = 0.14$, $t = -0.25$, $p = 0.8$), suggesting that statements made by L2 adult speakers were not rated more or less credible than statements made by L1 child speakers (see Table 1).

Results of Experiment 1 showed no significant difference in credibility ratings between L2 adult statements and L1 child statements (L2 adult: raw mean = 5.9, SD = 3.7; L1 child: raw mean = 5.9, SD = 3.8) (see Figure 1). Previous literature has found more positive attitudes toward L1 children than L1 adults, and also more negative attitudes toward L2 adults than L1 adults (Dixon et al., 2002). In a comparison of L2 adults with L1 children, social attitudes toward L1 children can therefore most certainly be expected to be clearly more positive than toward L2 adults, and an influence of social attitudes on credibility ratings should have favored L1 child statements over L2 adult statements. Predictions for an influence of perceptual fluency have been less clear. Speakers of both groups could be expected to be easily recognizable as members of their group, and both groups are known to be more variable in their productions than L1 adult speakers, which should decrease perceptual fluency for their recordings. It is, however, unclear if such a decrease was comparably large for both speaker groups in Experiment 1. The longer duration of the recordings for the L2 adult speakers might indicate otherwise, for example. *Post-hoc* interpretation showed that the current results are fully in

2 Note that Lev-Ari and Keysar (2010) had not found any evidence for an effect of *knowledge*. Identical to their statistical analysis, we additionally tested all statements and included the interaction of *speaker* (i.e., adults, children) and *knowledge* (i.e., yes, no, unsure) to the model. Similarly to Lev-Ari and Keysar (2010), *knowledge* did not improve the model.

3 In a *post-hoc* analysis, we also included *order* to the model. It did not change the results (*speaker*: $b = -0.06$, SD = 0.16, $t = -0.4$, $p = 0.68$; *order*: $b = 0.01$, SD = 0.01, $t = 1.16$, $p = 0.2$). While we did not collect data on familiarity with child speech, we did collect data regarding participants' prior experience with Arabic, Chinese, Spanish, and Russian. However, participants' familiarity with the foreign accents did not have an effect on *speaker* ($b = -0.04$, $SE = 0.15$, $t = -0.24$, $p = 0.8$).

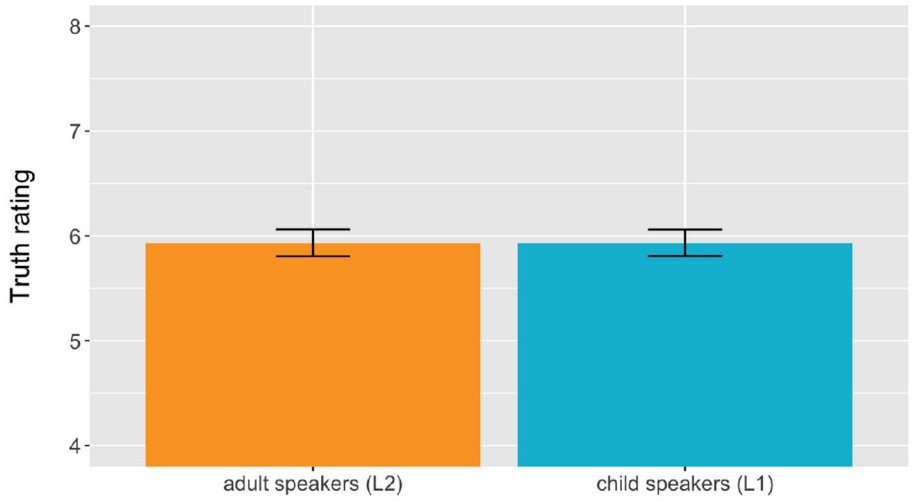

FIGURE 1
Average credibility ratings as a function of speaker (L2 adults, L1 children) of the raw data in Experiment 1. Higher numbers on the *y*-axis indicate higher perceived credibility. On the scale participants had used for their ratings, 0 had equaled "definitely false" and 14 "definitely true." The vertical bars represent standard errors.

line with an account based on pragmatic lenience. Both L2 adult speakers and L1 child speakers have a lower linguistic competence than L1 adult speakers, and pragmatic lenience advocates that participants' beliefs about lower competence of a speaker lead to more favorable ratings. In Experiment 2, we compared the credibility ratings for newly recorded L1 adult speakers (i.e., speakers with a high linguistic competence) with those for the L1 child speakers of Experiment 1 (i.e., speakers with a low linguistic competence).

3 Experiment 2

3.1 Method

3.1.1 Participants

Twenty-seven native listeners of German, between 20 and 33 years old (mean age = 25; SD = 3.4), participated in the experiment for a small monetary compensation. All participants were female students at the University of Tübingen and had no reported visual or hearing impairments.

3.1.2 Materials

The trivia statements and the recordings of the four L1 child speakers were the same ones used in Experiment 1. Recordings from the four L2 adult speakers in Experiment 1, were substituted with recordings from four L1 German adult speakers (mean age = 40, age range 37–47), who were (like the L1 child speakers) living in the Tübingen area at the time of the recordings. Care was taken again, that the recordings were fluent and free of hesitations. The style of speaking was natural with no noticeable hyper- or hypoarticulation. L1 adult speakers had an average F0 value of 201 Hz, in comparison to 256.8 Hz for the L1 child speakers. The F0 difference between the two groups of speakers (i.e., L1 adults vs. L1 children) was again significant ($t = -25.3$, $df = 358$, Cohen's $d = 2.6$, $p < .001$). Statements made by the L1 children did not differ reliably in duration from statements made by the L1 adult speakers (L1 children: 3,389.6 ms, L1 adults: 3,233.6 ms; $t = 1.63$, $df = 358$, Cohen's $d = 0.17$, $p = 0.1$).

3.1.3 Procedure

Experiment 2 was conducted in-person at the *LingTülab* of the university of T'ubingen. The experiment was controlled with Excel Visual Basic (version 16.0.11328.20362). Participants wore over-the-ear headphones and were tested individually. The experimental procedure was furthermore identical to Experiment 1.

3.2 Results

R (R Core Team, 2021, version 3.5.0) and lme4 (Bates et al., 2015) were used to perform linear mixed effect analyses on listeners' perceived credibility ratings. Similar to Experiment 1, only statements that were unknown to the participants were analyzed (70.6% of the data).[4] Participants had indicated for every statement that they had understood it and no further data points were removed from the analysis.

Similar to Experiment 1, the model included *ratings* as the dependent variable, *speaker* (L1 adult and L1 child, sum coded as 0.5 and −0.5, respectively) as independent variable and the random effect structure comprised *by-item* varying intercepts, and *by-participant* varying intercepts and varying slopes for *speaker*.[5] The results showed a significant effect for speaker ($b = -0.37$, $SE =$

4 Identical to the statistical analysis of Experiment 1 and Lev-Ari and Keysar (2010), we additionally tested all statements and included the interaction of *speaker* (i.e., L1 adults, L1 children) and *knowledge* to the model. It did not enhance the model and the pattern of results did not significantly change.

TABLE 2 LMER result summary for Experiment 2.

Fixed effects	*b*	SE	*t*	*Pr*>|*t*|	
(Intercept)	6.68	0.32	20.84	<2e-16	***
Speaker	−0.37	0.18	−1.96	0.047	*
Random effects	**Variance**	**SD**			
Participants	0.94	0.97			
Items	2.62	1.62			

*p<.05, ***p<.001.

0.18, t = −1.96, p = .04), such that listeners found trivia statements spoken by the L1 children to be more credible than statements made by the L1 adult speakers (see Table 2).

Experiment 2 found that L1 adult listeners rated trivia statements made by L1 child speakers as more credible than statements made by L1 adult speakers (L1 adult: raw mean = 6.3, SD = 3.3, L1 child: raw mean = 6.7, SD = 3.3) (see Figure 2). More favorable credibility ratings for L1 children are, for example, in line with an account based on social attitudes, since attitudes toward L1 children can be expected to be more positive in comparison to attitudes toward L1 adults (Axt et al., 2014; Chopik and Giasson, 2017). In terms of perceptual fluency, it seems unlikely that recordings from the L1 adults were more demanding to process than recordings from the L1 children, and a positive bias in the ratings toward L1 children is therefore unlikely to reflect a perceptual fluency disadvantage for the L1 adult speakers. The account that can best explain the pattern of results of both Experiments 1 and 2, is based on pragmatic lenience (Fairchild and Papafragou, 2018; Lorenzoni et al., 2022). Both L2 adults and L1 children have a lower linguistic competence than L1 adults. In this sense, both groups of speakers could have benefited from pragmatic lenience in Experiment 1 with no significant difference in credibility ratings, while in Experiment 2, pragmatic lenience favored the L1 child speakers over the L1 adult speakers.

4 General discussion

The purpose of the present study was to examine the effects of nativeness and speaker age on credibility in a German context. We adopted the methodology from Lev-Ari and Keysar (2010), who investigated ratings for L2 adults and L1 adult speakers, but compared trivia statements such as "*Das Auge des Vogel Strauss ist grösser als sein Gehirn*," "an ostrich's eye is bigger than its brain," spoken by L2 adults vs. L1 children (Experiment 1) and L1 adults vs. L1 children (Experiment 2). Overall, findings showed a diverging pattern of credibility judgments across Experiments. While the results of Experiment 1 showed neither higher nor lower credibility ratings for L2 adult speakers in comparison to L1 child speakers, findings of Experiment 2 demonstrated a difference in credibility ratings for speaker groups, such that statements made by L1 child speakers were rated as more credible than statements made by L1 adult speakers. Thus in contrast to Lev-Ari and Keysar (2010) and Boduch-Grabka and Lev-Ari (2021), statements made by native adult speakers had not received more positive ratings in the present study.

In Experiment 1, no significant effect of speaker group was found, and an interpretation of a null effect must of course be taken with caution. However, given that more positive attitudes have previously been attested toward children than adults (Dixon et al., 2002), an influence of social attitudes should have favored the trivia statements made by the L1 children in Experiment 1. This was not found. At first glance, results of Experiment 1 could rather imply that the speech signal had been harder to process for both speakers groups, i.e., that perceptual fluency had been reduced for L2 adult speakers and L1 child speakers alike, affecting credibility ratings to an equal extent. We cannot rule out this possibility, but we would like to note that L1 children and L2 adults showed some distinct variations in their productions. While it can be expected, based on previous literature (Vorperian and Kent, 2007; Wade et al., 2007), that both L2 adults and L1 children exhibited some variability in their pronunciation related to the ongoing process of language learning, L1 children additionally differed in the height of their voices (F0) from L2 adults. The size of the vocal tract, including the pharyngeal and oral cavities, can influence the characteristics of the voice. Children have a smaller vocal tract and oral cavities, which can modulate both the fundamental frequency (F0) and the spectral quality of their speech signal, i.e. raises the resonating frequencies (formants) (e.g., Peterson and Barney, 1952; Hillenbrand et al., 1995; Vorperian and Kent, 2007). Furthermore, an impressionistic analysis of the recordings by the authors suggested that the L2 adult speakers possibly had made more segmental substitutions than the children had. The recordings of the L2 adult speakers had also been significantly longer than the recordings from the children. While it is difficult to draw firm conclusions from this about the ease with which the recordings had been processed (i.e, perceptual fluency), we would like to point out that care was taken that all recordings had been intelligible in the sense that the listeners understood what had been said, so that the ratings did not reflect a lack of understanding. Boduch-Grabka and Lev-Ari (2021) argued furthermore in 2021 for an influence of familiarity, since brief exposure to a foreign accent reduced the negative bias in credibility ratings due to an improvement in the processing of the accent. Contrary to Boduch-Grabka and Lev-Ari (2021) who found that with little pre-exposure to an accent subsequent credibility ratings increased, we found no evidence that exposure during the rating task resulted in higher ratings, neither in Experiment 1 nor in Experiment 2.

5 In a *post-hoc* analysis, we also included *order* to the model. The pattern of the results did not change (*speaker*: b = −0.03, SD = 0.18, t = −2.0, p = 0.04); *order* (b = 0.005, SD = 0.005, t = 0.99, p = 0.32).

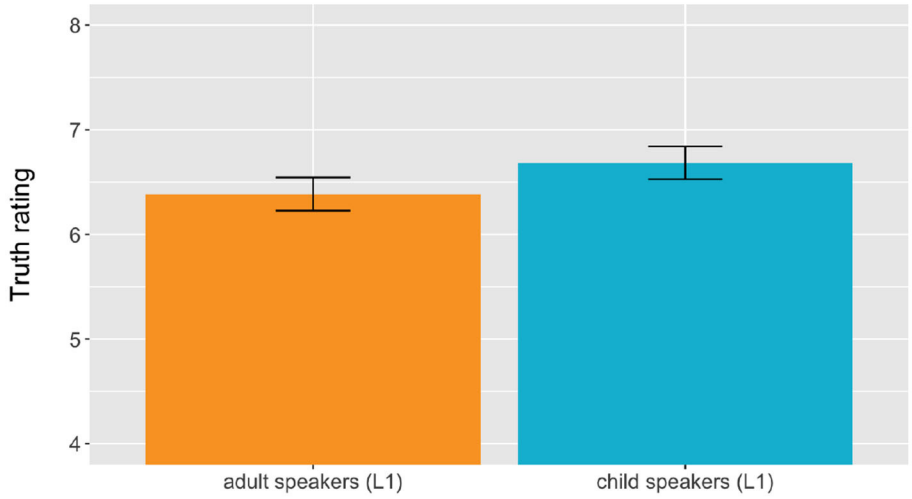

FIGURE 2
Average credibility ratings as a function of speaker (L1 adults, L1 children) of the raw data in Experiment 2. Higher numbers on the *y*-axis indicate higher perceived credibility. On the scale participants had used for their ratings, 0 had equaled "definitely false" and 14 "definitely true." The vertical bars represent standard errors.

Specifically, we found no influence of order on the credibility ratings. The third explanation that has been put forward in the literature for biases in credibility ratings is based on pragmatic lenience. Particularly, Lorenzoni et al. (2022) recently found a positive bias for credibility ratings for written trivia statements that could be attributed to non-native speakers rather than to native speakers. Since perceptual fluency could not directly account for their findings (the statements were presented in written form), the authors argued for an effect of pragmatic lenience toward non-native speakers that is related to participants' beliefs about these speakers' lower linguistic competence. If such pragmatic lenience is extended to children, who also have a lower linguistic competence, then the lack of a difference in Experiment would be in line with an account of pragmatic lenience.

In Experiment 2, the experimental design was identical to Experiment 1, but the recordings of the L2 adult speakers were substituted with recordings from L1 adult speakers. This time, statements produced by the L1 child speakers received significantly higher credibility ratings than statements produced by the L1 adult speakers. While initially this might seem surprising, considering that children have less world knowledge than adults, such a positive bias for statements made by L1 children is in line with accounts of social attitudes and pragmatic lenience (Chopik and Giasson, 2017; Fairchild and Papafragou, 2018; Lorenzoni et al., 2022). Both accounts would predict higher credibility ratings for statements made by children, the former due to more positive attitudes toward children and the latter due to pragmatic lenience toward speakers with a lower linguistic competence. An interpretation of the results in terms of perceptual fluency would require that the recordings from the L1 adult speakers had been harder to process than the recordings from the L1 children (e.g., Lev-Ari and Keysar, 2010). Since all recordings had been fluent, the duration of the statements had been neither shorter nor longer for the L1 adult speakers in comparison to the statements of the L1 child speakers, and neither speaker group had noticeably hyper- or hypoarticulated their productions, reduced perceptual fluency for the L1 adult speakers seems doubtful. As such, perceptual fluency is an unlikely account for the results in Experiment 2. While the findings of Experiment 2 are in line with both an account of social attitudes and pragmatic lenience, only pragmatic lenience can explain the results of both Experiments 1 and 2, since lower linguistic competence should result in lenience toward statements from both L2 adults and L1 children (Experiment 1) and result in more lenience toward statements from L1 children in comparison to L1 adults (Experiment 2).

The study by Lorenzoni et al. (2022) compared credibility ratings for L2 adult statements with that for L1 adult statements and found a positive bias for L2 adult statements, explained with an account of pragmatic lenience (see also, Fairchild and Papafragou, 2018). We did not directly compare L2 adult statements with L1 adult statements in the present study and therefore could not replicate their effects, but by extension the positive bias toward L1 children in Experiment 2 (in comparison with L1 adults), and the lack of a difference in Experiment 1 (in comparison with L2 adults) implies that not only statements made by L1 children but also statements made by L2 adult speakers were rated more leniently. This is of course good news on a more general level, as it suggests that our evolving multicultural society may still be biased in their social evaluation of various speaker groups, but utterances from speakers with lower linguistic competence are not necessarily met with negative evaluations but rather with forgiveness.

While most previous studies investigated credibility judgments made by native adult listeners (e.g., Souza and Markman, 2013; Podlipský et al., 2016; Stocker, 2017; Frances et al., 2018), a few studies have also looked at other participant groups such as non-native listeners (Hanzlíková and Skarnitzl, 2017) and children (Avarino et al., 2021). Taken together, credibility ratings have been collected now in contexts that vary in language of the statements, participants' task, accent strength, modality, and linguistic competence and age of both the speakers of the

statements and the participants. The emerging differences in results highlight the importance of the various factors that can influence credibility judgments, and these differences can maybe partially be explained in terms of differences in context. For example, while the majority of previous studies have attested negative attitudes toward foreign-accented speakers, a few have also found positive reactions (Dewaele and McCloskey, 2015). A Spanish accent is considered a nonstandard accent in the United States and is connected with negative traits, but it has been shown to positively affect listeners' perception of speakers' educational background, social status, and personal traits like attractiveness in the United Kingdom (Fuertes et al., 2012). Negative attitudes toward L2 speakers can thus vary across contexts. Similarly, foreign accent strength and familiarity can vary and have been found to co-determine perceptual fluency for native listeners, such that words with the same accent markers are processed differently by native listeners that vary in experience with the accent (Witteman et al., 2013), and acoustic similarity and perceived accentedness are not always predictive of processing difficulties (Witteman et al., 2011).

In conclusion, the present results highlight the important role of speaker in credibility ratings such that distinct patterns of credibility can emerge in different speaker contexts. The results of Experiment 1 found no evidence that statements made by L2 adult speakers are perceived as either less or more credible than statements from L1 child speakers. In combination with Experiment 2, in which statements made by L1 children were rated as more credible than statements from L1 adults, the results are in line with an account of pragmatic lenience that advocates lenience toward speakers with lower linguistic competence which in turn mediates credibility ratings.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Prof. Dr. Petra Schumacher, DGFS (Deutsche Gesellschaft für Sprachwissenschaft). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

TLT: Data curation, Formal analysis, Investigation, Methodology, Project administration, Visualization, Writing—original draft. AW: Conceptualization, Supervision, Writing—review & editing, Project administration.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. We acknowledge support by the Open Access Publishing Fund of the University of Tübingen.

Acknowledgments

We would like to thank the LingTüLab for support with laboratory space and student assistants' help with data collection.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Anderson, S., Downs, S. D., Faucette, K., Griffin, J., King, T., Woolstenhulme, S., et al. (2007). How accents affect perception of intelligence, physical attractiveness, and trustworthiness of Middle-Easter-, Latin-American-, British- and Standard-American-English-accented speakers. *Intuition: BYU Undergrad. J. Res.* 3, 5–11.

Anwyl-Irvine, A., Massonnié, J., Flitton, A., Kirkham, N., and Evershed, J. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behav. Res. Methods* 52, 388–407. doi: 10.3758/s13428-019-01237-x

Avarino, A., Thorburn, R., and White, K. (2021). "Broken telephone: children's judgments of messages delivered by non-native speakers are influenced by processing fluency," in *Proceedings of the Annual Meeting of the Cognitive Science Society, 43*, 2349–2355.

Axt, J. R., Ebersole, C. R., and Nosek, B. A. (2014). The rules of implicit evaluation by race, religion, and age. *Psychol. Sci.* 25, 1804–1815. doi: 10.1177/0956797614543801

Bala, N., Ramakrishnan, K., Lindsay, R., and Lee, K. (2005). Judicial assessment of the credibility of child witnesses. *Alta. Law Rev.* 42, 995–1017. doi: 10.29173/alr1270

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). lme4: linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Baus, C., McAleer, P., Marcoux, K., Belin, P., and Costa, A. (2019). Forming social impressions from voices in native and foreign languages. *Sci. Rep.* 9, 414. doi: 10.1038/s41598-018-36518-6

Boduch-Grabka, K., and Lev-Ari, S. (2021). Exposing individuals to foreign accent increases their trust in what nonnative speakers say. *Cogn. Sci.* 45, e13064. doi: 10.1111/cogs.13064

Boersma, P., and Weenink, D. (2023). *Praat: Doing Phonetics by Computer [computer Program], version 6.3.09.*

Bürki-Cohen, J., Miller, J. L., and Eimas, P. D. (2001). Perceiving non-native speech. *Lang. Speech* 44, 149–169. doi: 10.1177/00238309010440020201

Chopik, W. J., and Giasson, H. L. (2017). Age differences in explicit and implicit age attitudes across the life span. *Gerontologist* 57(suppl_2), S169–S177. doi: 10.1093/geront/gnx058

Cooper, A., Fecher, N., and Johnson, E. K. (2020). Identifying children's voices. *J. Acoust. Soc. Am.* 148, 324–333. doi: 10.1121/10.0001576

Cristia, A., Seidl, A., Vaughn, C., Schmale, R., Bradlow, A., Floccia, C., et al. (2012). Linguistic processing of accented speech across the lifespan. *Front. Psychol.* 3, 479. doi: 10.3389/fpsyg.2012.00479

De Meo, A. (2012). "How credible is a non-native speaker? Prosody and surroundings," in *Methodological Perspectives on Second Language Prosody. Papers from ML2P* (Padova), 3–9.

Dewaele, J.-M., and McCloskey, J. (2015). Attitudes towards foreign accents among adult multilingual language users. *J. Multiling. Multicult. Dev.* 36, 221–238. doi: 10.1080/01434632.2014.909445

Dixon, J., Mahoney, B., and Cocks, R. (2002). Accents of guilt? Effects of regional accent, race, and crime type on attributions of guilt. *J. Lang. Soc. Psychol.* 21, 162–168. doi: 10.1177/02627X02021002004

Dragojevic, M., and Giles, H. (2016). I don't like you because you're hard to understand: the role of processing fluency in the language attitudes process. *Hum. Commun. Res.* 42, 396–420. doi: 10.1111/hcre.12079

Dragojevic, M., Giles, H., Beck, A.-C., and Tatum, N. (2017). The fluency principle: why foreign accent strength negatively biases language attitudes. *Commun. Monogr.* 84, 1–21. doi: 10.1080/03637751.2017.1322213

Edwards, J. (1999). Refining our understanding of language attitudes. *J. Lang. Soc. Psychol.* 18, 101–110. doi: 10.1177/0261927X99018001007

Fairchild, S., and Papafragou, A. (2018). Sins of omission are more likely to be forgiven in non-native speakers. *Cognition* 181, 80–92. doi: 10.1016/j.cognition.2018.08.010

Flege, J. E. (1984). The detection of French accent by American listeners. *J. Acoust. Soc. Am.* 76, 692–707. doi: 10.1121/1.391256

Foucart, A., Costa, A., Morís-Fernndez, L., and Hartsuiker, R. J. (2020). Foreignness or processing fluency? On understanding the negative bias toward foreign-accented speakers. *Lang. Learn.* 70, 974–1016. doi: 10.1111/lang.12413

Foucart, A., and Hartsuiker, R. J. (2021). Are foreign-accented speakers that "incredible"? The impact of the speaker's indexical properties on sentence processing. *Neuropsychologia* 158, 107902. doi: 10.1016/j.neuropsychologia.2021.107902

Foucart, A., Santamar-a-Garc-a, H., and Hartsuiker, R. J. (2019). Short exposure to a foreign accent impacts subsequent cognitive processes. *Neuropsychologia* 129, 1–9. doi: 10.1016/j.neuropsychologia.2019.02.021

Frances, C., Costa, A., and Baus, C. (2018). On the effects of regional accents on memory and credibility. *Acta Psychol.* 186, 63–70. doi: 10.1016/j.actpsy.2018.04.003

Frumkin, L. A., and Stone, A. (2020). Not all eyewitnesses are equal: accent status, race and age interact to influence evaluations of testimony. *J. Ethn. Crim. Justice* 18, 123–145. doi: 10.1080/15377938.2020.1727806

Fuertes, J. N., Gottdiener, W., Martin, H., Gilbert, T., and Giles, H. (2012). A meta analysis of the effects of speakers' accents on interpersonal evaluations. *Eur. J. Soc. Psychol.* 42, 120–133. doi: 10.1002/ejsp.862

Fuertes, J. N., Potere, J. C., and Ramirez, K. Y. (2002). Effects of speech accents on interpersonal evaluations: Implications for counseling practice and research. *Cult. Divers. Ethn. Minor. Psychol.* 8, 346–56. doi: 10.1037/1099-9809.8.4.347

Gibson, E., Tan, C., Futrell, R., Mahowald, K., Konieczny, L., Hemforth, B., et al. (2017). Don't underestimate the benefits of being misunderstood. *Psychol. Sci.* 28, 703–712. doi: 10.1177/0956797617690277

Gluszek, A., and Dovidio, J. F. (2010). The way they speak: a social psychological perspective on the stigma of nonnative accents in communication. *Pers. Soc. Psychol. Rev.* 14, 214–237. doi: 10.1177/1088868309359288

Hanzlíková, D., and Skarnitzl, R. (2017). Credibility of native and non-native speakers of English revisited: do non-native listeners feel the same? *Res. Lang.* 15, 285–298. doi: 10.1515/rela-2017-0016

Hillenbrand, J. M., Getty, L., Clark, M. J., and Wheeler, K. (1995). Acoustic characteristics of American English vowels. *J. Acoust. Soc. Am.* 97, 3099–3111. doi: 10.1121/1.411872

Kent, R. D. (1976). Anatomical and neuromuscular maturation of the speech mechanism: evidence from acoustic studies. *J. Speech Hear. Res.* 19, 421–47. doi: 10.1044/jshr.1903.421

Kreiman, J., and Sidtis, D. (2011). *Foundations of Voice Studies: An Interdisciplinary Approach to Voice Production and Perception*. Hoboken, NJ: Wiley-Blackwell Publication, John Wiley and Sons. doi: 10.1002/9781444395068

Latinus, M., and Belin, P. (2011). Human voice perception. *Curr. Biol.* 21, R143–R145. doi: 10.1016/j.cub.2010.12.033

Lee, S., Potamianos, A., and Narayanan, S. (1999). Acoustics of children's speech: developmental changes of temporal and spectral parameters. *J. Acoust. Soc. Am.* 105, 1455–1468. doi: 10.1121/1.426686

Lev-Ari, S., and Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *J. Exp. Soc. Psychol.*, 46, 1093–1096. doi: 10.1016/j.jesp.2010.05.025

Lindemann, S. (2003). Koreans, Chinese or Indians? Attitudes and ideologies about non-native English speakers in the United States. *J. Socioling.* 7, 348–364. doi: 10.1111/1467-9481.00228

Lorenzoni, A., Pagliarini, E., Vespignani, F., and Navarrete, E. (2022). Pragmatic and knowledge range lenience towards foreigners. *Acta Psychol.* 226, 103572. doi: 10.1016/j.actpsy.2022.103572

Milne, A. E., Bianco, R., Poole, K. C., Zhao, S., Oxenham, A. J., Billig, A. J., et al. (2021). An online headphone screening test based on dichotic pitch. *Behav. Res.* 53, 1551–1562. doi: 10.3758/s13428-020-01514-0

Munro, M. J., Derwing, T. M., and Burgess, C. S. (2010). Detection of nonnative speaker status from content-masked speech. *Speech Commun.* 52, 626–637. doi: 10.1016/j.specom.2010.02.013

Munro, M. J., Derwing, T. M., and Sat, K. (2006). Salient accents, covert attitudes: Consciousness-raising for pre-service second language teachers. *Prospect* 21, 67–79.

Oppenheimer, D. M. (2008). The secret life of fluency. *Trends Cogn. Sci.* 12, 237–241. doi: 10.1016/j.tics.2008.02.014

Peterson, G. E., and Barney, H. L. (1952). Control methods used in a study of the vowels. *J. Acoust. Soc. Am.* 24, 175–184. doi: 10.1121/1.1906875

Pietraszewski, D., and Schwartz, A. (2013). Evidence that accent is a dimension of social categorization, not a byproduct of perceptual salience, familiarity, or ease-of-processing. *Evol. Hum. Behav.* 35, 43–50. doi: 10.1016/j.evolhumbehav.2013.09.006

Podlipský, V., Šimáčková, R., and Petráž, D. (2016). Is there an interlanguage speech credibility benefit? *Top. Linguist.* 17, 30–44. doi: 10.1515/topling-2016-0003

Potamianos, A., and Narayanan, S. (2003). Robust recognition of children's speech. *IEEE Trans. Speech Audio Process.* 11, 603–616. doi: 10.1109/TSA.2003.818026

R Core Team (2021). *R: A Language and Environment for Statistical Computing. Versions 3.5.0 to 4.0.5*. Vienna: R Foundation for Statistical Computing.

Reber, R., Schwarz, N., and Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Pers. Soc. Psychol. Rev.* 8, 364–382. doi: 10.1207/s15327957pspr0804_3

Romero-Rivas, C., Morgan, C., and Collier, T. (2021). Accentism on trial: categorization/stereotyping and implicit biases predict harsher sentences for foreign-accented defendants. *J. Lang. Soc. Psychol.* 41, 191–208. doi: 10.1177/0261927X211022785

Ross, D. F., Dunning, D., Toglia, M. P., and Ceci, S. J. (1990). The child in the eyes of the jury: assessing mock jurors' perceptions of the child witness. *Law Hum. Behav.* 14, 5–23. doi: 10.1007/BF01055786

Schwarz, N. (2004). Meta-cognitive experiences in consumer judgment and decision making. *J. Consum. Psychol.* 14, 332–348. doi: 10.1207/s15327663jcp1404_2

Seggie, I. (1983). Attribution of guilt as a function of ethnic accent and type of crime. *J. Multiling. Multicult. Dev.* 4, 197–206. doi: 10.1080/01434632.1983.9994111

Smith, B., Sugarman, M., and Long, S. (1983). Experimental manipulation of speaking rate for studying temporal variability in children's speech. *J. Acoust. Soc. Am.* 74, 744–749. doi: 10.1121/1.389860

Souza, A. L., and Markman, A. B. (2013). "Foreign accent does not influence cognitive judgments," in *Proceedings of the 35th Annual Meeting of the Cognitive Science Society* (Berlin), 35, 1360–1365.

Stocker, L. (2017). The impact of foreign accent on credibility: an analysis of cognitive statement ratings in a Swiss context. *J. Psycholinguist. Res.* 46, 617–628. doi: 10.1007/s10936-016-9455-x

Tingley, B. M., and Allen, G. D. (1975). Development of speech timing control in children. *Child Dev.* 46, 186–194. doi: 10.2307/1128847

Truong, T. L., and Weber, A. (2020). "Trust issues: the effect of speaker age on credibility," in *Proceedings of Linguistic Evidence 2020: Linguistic Theory Enriched by Experimental Data* (Tübingen), 351–361.

Vorperian, H. K., and Kent, R. D. (2007). Vowel acoustic space development in children: a synthesis of acoustic and anatomic data. *J. Speech Lang. Hear. Res.* 50, 1510–45. doi: 10.1044/1092-4388(2007/104)

Wade, T., Jongman, A., and Sereno, J. (2007). Effects of acoustic variability in the perceptual learning of non-native-accented speech sounds. *Phonetica* 64, 122–144. doi: 10.1159/000107913

Witteman, M. J., Weber, A., and McQueen, J. M. (2011). "On the relationship between perceived accentedness, acoustic similarity, and processing difficulty in foreign-accented speech," in *Proceedings of the Annual Conference of the International Speech Communication Association* (Florence), 2229–2232. doi: 10.21437/Interspeech.2011-583

Witteman, M. J., Weber, A., and McQueen, J. M. (2013). Foreign accent strength and listener familiarity with an accent co-determine speed of perceptual adaptation. *Attent. Percept. Psychophys.* 75, 537–556. doi: 10.3758/s13414-012-0404-y