Same as always: I suck at titles

Esperanza Badaya[1], Robert J. Hartsuiker[1], and & Martin Corley[2]

[1] Ghent University

[2] University of Edinburgh

Author Note

Correspondence concerning this article should be addressed to Esperanza Badaya. E-mail: esperanza.badaya@ugent.be

Abstract

Whether an individual is perceived as knowledgeable by others can be biased by several potentially irrelevant factors, ranging from the ephemeral (e.g., how speech is produced) to the situational (e.g., who is speaking). Existing evidence for these effects tends to come from metalinguistic evaluations; here we examine the behavioral consequences of variations in the ways in which spoken messages are produced. Specifically,we investigate how attributions of knowledgeability are influenced by speech fluency, and by whether or not the speaker has a native accent, cues which have been previously shown to bias listeners' evaluations of speakers and what they are saying. We use a novel horse-race betting paradigm, in which participants have to place bets on horses which have been described in short spoken passages. By manipulating the fluency (fluent or disfluent) and accent (native or non-native) with which the descriptions are produced, we show that listeners are less likely to follow advice from a speaker who is disfluent, regardless of whether they are disfluent for reasons other than low knowledge (e.g., nativeness). Overall, our results align with a broader body of literature suggesting that the effects of hesitation phenomena are context-independent. Importantly, we demonstrate that individuals' meta-linguistic judgements can be measured through behavior, suggesting that these judgements have ecological importance beyond the laboratory.

Keywords: filled pauses, disfluencies, non-native-accented speech, feeling of another's knowing

Word count: 5367

Same as always: I suck at titles

Impressions of other people arise naturally and automatically (Uleman et al., 2008). Speakers can be evaluated not only by what they say but also by how they say it. For example, features of speech can affect perceptions of confidence and persuasion (Guyer et al., 2019), status and solidarity (Pittam & Gallois, 1986), or even attractiveness (Feinberg et al., 2005). Similarly, speech which includes disfluencies, such as *uh* or *um* in English, can lead to poorer evaluations of the speaker in terms of intelligence (Christenfeld, 1995), competence (Norton-Ford & Hogan, 1980), and certainty of their own knowledge (Brennan & Williams, 1995). Interestingly, speech produced with an accent different from one's own—particularly a *foreign* accent—can also elicit negative perceptions of the speaker, including worse evaluations in terms of status, solidarity, and credibility (Gluszek et al., 2011; Gluszek & Dovidio, 2010; Lev-Ari & Keysar, 2010; Rakić et al., 2011). Most of these studies, however, have relied on listeners' explicit judgements, which may not reflect everyday language comprehension (Armstrong et al., 1983): Do these judgements arise even when individuals are not asked to make them?. In this study we tested whether listeners' evaluations of a speaker affect their decision-making, and if the co-presence of these two factors (i.e., (dis)fluency and foreign accents) may diminish the negative impact they have individually. In what follows, we briefly describe how speech fluency and accent have been shown to bias listeners' evaluations.

There is a growing body of literature showing that certain speech features impact listeners' perceptions of a speaker's knowledgeability. Voices with a slow speech rate, low amplitude, and a large F0 range are less likely to be rated as confident, and are associated with distinct neural responses in the listener (Jiang & Pell, 2015, 2016b, 2016a). Similarly, prosody guides listeners' evaluations of certainty and honesty cross-linguistically (Goupil et al., 2021). These non-verbal qualities of speech may be accompanied by verbal markers of hesitation, such as disfluencies (Grosjean & Deschamps, 1975; Jiang & Pell, 2017; Shriberg et

al., 1997; Shriberg & Lickley, 1993). Brennan and Williams (1995) demonstrated that filled pauses (e.g., uh, mm, or um) impact the perception of speakers' confidence in their knowledge. In their study, participants listened to previously recorded answers to trivia questions (without hearing the questions) and were asked to rate how likely each speaker would be to recognise the correct answer to the question i.e., to rate their *feeling of another's knowing*, FOAK. Among the cues that biased participants' assessments were filled pauses: Answers containing a disfluency were more likely to receive lower FOAK ratings, suggesting that filled pauses were taken as reflective of speakers' reduced certainty about their knowledge. In contrast, non-answers (i.e., 'I don't know') were more likely to receive a higher FOAK rating if preceded by a filled pause. Brennan and Williams took this as evidence that listeners are sensitive to the surface form of delivery, and in particular, to the cues displayed by speakers when they do not know (or cannot remember at the moment of being asked) the answer to a question (see also Smith & Clark, 1993).

In a similar vein, speaker identity can itself affect evaluations of both the speaker and the content of speech. In particular, a speaker's nativeness (whether they are producing speech in their first or second language) has been consistently reported to elicit different evaluations. On the one hand, speakers with a foreign accent are more likely to be negatively evaluated (e.g., Dragojevic & Giles, 2016; Gluszek & Dovidio, 2010) and statements produced with a foreign accent are more likely to receive lower ratings of credibility compared to statements produced with a native accent (Lev-Ari & Keysar, 2010; Boduch-Graba & Lev-Ari, 2021; but cf. Souza & Markan, 2013; Stocker, 2017; Foucart, Costa, Morís-Fernández, & Hartsuiker, 2020; Foucart & Hartsuiker, 2021; Wetzel, Zufferey, & Gygax, 2021; Barlow et al., 2024). These negative effects triggered by a speaker's non-nativeness have been primarily accounted for in terms of processing fluency (Lev-Ari & Keysar, 2010), whereby difficulties in understanding accented speech decreases processing fluency and consequently elicits negative evaluations about the source.

On the other hand, there seem to be instances where being a non-native speaker is advantageous. Fairchild et al. (2020) examined whether and how failures to provide all relevant information (i.e., being under-informative) are interpreted differently as a function of speakers' nativeness. In four experiments, participants read stories where a speaker (native/non-native) provided a description. Crucially, these descriptions failed to provide all relevant information (e.g., not naming all elements on a display, failing to list all uses of an invention). Participants were then asked to provide an explanation for the speaker's under-informativeness. As descriptions were presented in the written modality, any differences in attributions of under-informativeness would be solely guided by the putative speakers' identities, rather than, for example, by processing fluency. Native speakers who were under-informative were more likely to be rated as unwilling to share the information than under-informative non-native speakers. The same pragmatic failure by a non-native speaker, however, was taken as a sign of inability to produce the necessary information (Exp. 1), even when participants were not explicitly informed that the non-native speaker could experience language difficulties (Exp. 2). Further, this 'forgiveness' for underinformative statements had consequences for participants' subsequent behaviours: They were more likely to learn new information from a previously encountered underinformative non-native speaker than from a native speaker (Exp. 3 and 4) (for similar findings, see Fairchild & Papafragou, 2018; Lorenzoni et al., 2022). Interestingly, this pattern of results holds even for auditory stimuli (Ip & Papafragou, 2022), suggesting that difficulties in comprehending non-native-accented speech alone do not necessarily lead to negative evaluations of the speaker.

One potential explanation for Fairchild et al.'s (2020) results is that non-native speakers are evaluated differently from native speakers due to expectations about their competence. Non-native speakers' accents may invoke stereotypes that suggests that these speakers are less competent linguistically than native speakers (Fairchild & Papafragou, 2018; Lev-Ari, 2015). This expectation-based account proposes that stereotypes of non-native

speakers' low linguistic competence affect the ways in which their speech is comprehended and interpreted. For example, in the case of under-informativeness, signs that can be attributed to low linguistic competence would be interpreted as such when produced by a non-native speaker and be attributed to other factors, such as ill will, when produced by a native speaker. In line with this hypothesis, a speaker's nativeness affects how ironic a statement is perceived to be (Bazzi et al., 2022; Caffarra et al., 2018), or how syntactic errors are processed (Hanuíková et al., 2012).

Given that speakers producing speech in their second language are more disfluent (Bergmann et al., 2015; Gkalitsiou & Werle, 2023) and are perceived as such by native listeners (Pinget et al., 2014), it is possible that a speaker's non-nativeness leads to different interpretations of disfluencies. This would align with findings in the disfluency literature whereby the effects of disfluencies in speech comprehension are dependent on who produces them (Arnold et al., 2007; Barr & Seyfeddinipur, 2010; Heller et al., 2015), including non-native versus native speakers (Bosker, Quené, Sanders, & De Jong, 2014), allegedly because listeners are sensitive to the speaker's mental state and the reasons for being disfluent.

Recently, Matzinger et al. (2023) investigated whether listeners' perceptions of why a speaker was disfluent differed for native and non-native speakers. In their study, participants listened to staged conversations in which native and non-native speakers answered trivia questions and requests, and were explicitly asked to rate each speaker's knowledge and confidence (for FOAK) and their willingness to grant the request. Crucially, speakers' fluency was manipulated by having different inter-turn pauses: Answers were prefaced with either short (200 ms) or long (1200 ms) pauses. For requests, long pauses were more likely to be associated with unwillingness for native compared to non-native speakers. However, FOAK ratings did not differ between speakers: Long pauses produced by either speaker were likely to be taken as reflecting low confidence and low knowledge. Matzinger et al. attributed

this pattern to the degree with which a speaker's mental state is relevant across conversational contexts.

As is the case for Matzinger et al. (2023), most studies evaluating the effects of manner of speech on perceptions of the speaker have used explicit ratings. Participants are asked to rate particular traits of speakers or what they say on a scale. Explicit measures as those are more likely to be affected by factors such as self-presentation and less likely to reflect how individuals would behave (Greenwald et al., 2002). In contrast, implicit measures overcome these obstacles while still being able to predict behaviours: For example the Implicit Association Test can predict racist behaviours (McConnell & Leibold, 2001; although see Payne & Hannay, 2021 for whether IAT scores are better conceptualised as an aggregation rather than at the individual level). In the case of evaluations about the speaker, they tend to be elicited in non-social contexts: Participants are either overhearers (e.g., Matzinger et al., 2023) or simply attend to a previously recorded speaker (Brennan & Williams, 1995; Ip & Papafragou, 2022; e.g., Lev-Ari & Keysar, 2010). Further, in these experiments, participants have nothing at stake i.e., they are evaluating the speaker, but they do not have to perform any behaviour following the utterance they just heard. While experiments eliciting metalinguistic judgements shed light on what factors may affect how a speaker is perceived, the lack of interaction on participant's behalf entails that these findings may not explain how listeners evaluate speech implicitly, when to do so is consequential.

Here, we propose an implicit measurement of listeners' assessments of the speaker's certainty, using a horse-race paradigm. In this task, participants listen to a set of speakers provide descriptions of horses, and are asked to distribute virtual tokens as 'bets' on each horse's likelihood of winning a putative race. This approach presents two advantages over previous experiments. First, in the horse-race paradigm, participants are not explicitly asked to evaluate a specific trait of the speaker (in this case, how knowledgeable they are). Instead, we take participants' allocations of 'betting money' as an indirect measurement of their

perceptions of speakers' knowledge. Indeed, pilot studies have shown that individuals are sensitive to this manipulation and that disfluent information leads to smaller bets (Butterworth, 2019). Second, horse races provide a scenario where individuals can make decisions based on what they are told, but the content of speech itself may not be informative for many individuals (in that participants are not familiar with the world of horse racing or its technical vocabulary).

In a pre-registered study, we set up to test whether and how perceptions of knowledgeability are biased by manner of delivery in the form of fluency, and by the speaker's identity as conveyed by their accent. We presented participants with recordings of a native and a non-native speaker, each describing two horses, with one description produced fluently and one description produced disfluently. If listeners are sensitive to speaker identity when making judgements about knowledge and confidence following disfluent speech, then descriptions peppered with disfluencies provided by a native speaker should result in less money bet, reflecting listeners' lower trust in, or lower FOAK, for this speaker compared to disfluent descriptions provided by a non-native speaker. The non-native disfluent descriptions may not impact listeners' betting behaviours, to the extent that they consider the possibility of difficulties in production when assessing the speaker's knowledge. To further control for the potential effects of (non)-nativeness on certainty on its own, we measured participants' language attitudes towards each speaker (see Dragojevic & Giles, 2016), perceived fluency, accentedness, and comprehensibility of the native and the non-native speaker, as well as the perceived trustworthiness of each speaker. We additionally measured participants' familiarity with and exposure to native and non-native-accented English on a daily basis, to account for the fact that exposure to non-native accents can reduce their negative effects on listeners' judgements (Boduch-Grabka & Lev-Ari, 2021).

Methods

Our pre-registration, all experimental stimuli and scripts can be found at https://osf.io/zsut7/. Deviations from the original pre-registration (e.g., additional analysis) are marked as such.

Participants

We conducted a sensitivity power analysis via data simulation following DeBruine and Barr (2021). We explored the required number of participants for a study with a power higher than .8 for the interaction between fluency and speaker identity. We conducted 1000 simulations per different combinations of the effect size of the interaction (ranging from small to medium) and the standard deviation of the residuals. In this analysis, we assumed a medium effect size of fluency and no effect of speaker identity. This analysis showed that a sample size of 360 participants ensured enough power to detect a medium or greater effect size.

In our pre-registration, we stipulated that only participants born and raised, and currently residing in the United Kingdom, with English as their first and only language, and with no auditory disorders could take part in the study. Further, we would exclude from analyse data from participants who reported the experiment's aim or manipulation, rated the naturalness of the auditory stimuli (defined as how likely they believed the audios to have been recorded in one go) lower than four of a 9-point scale (see Procedure section), or considered themselves experts in horse races. This meant that we recruited 641 participants for a sample of 360 (N = 268 rated the audio lower than four, 9 reported the manipulation, 11 rated themselves as experts in horse races). Participants were aged between 18 and 30, and born and raised in the United Kingdom. Participants were recruited via the online platform Prolific, and gave their informed consent as approved by the University of Edinburgh PPLS Ethics Committee (ref no 133-2223/3). Participants were reimbursed £1.50 for a 10-minutes experiment.

Visual stimuli

We selected a candidate set of eight images of racehorses from the web. The selected images each featured only one racehorse in the foreground, in motion, ridden by a jockey, and the horses all took up approximately the same proportion of the image. To ensure that the pictures did not bias participants' bets, we recruited ten participants on Prolific, who did not take part in the main study, and asked them to rate on a 10-point scale how likely each horse was to win a hypothetical race and to rank the horses in the order they thought they would cross the finish line, in exchange for £0.45.

A one-way repeated measures ANOVA showed that there were no differences in how likely each horse was thought to be to win a race ($F(7, 72) = 0.50$, $p = .832$, $\hat{\eta}^2_G = .046$, 90% CI $[.000, .050]$). An ordinal logistic regression showed that none of the horses were more likely to be ranked differently from the others (all $|t| < 2$). We therefore selected four out of the eight images as visual stimuli.

Auditory stimuli

We used four descriptions of racing horses from *The Racing Post*, retrieved on October 2018, originally edited and used by Butterworth (2019). Each passage consisted of three to four sentences describing a horse and its performance in previous races. 30 British English speakers, who did not participate in the final experiment, rated these passages to ensure that all descriptions were perceived as equally likely to describe a winning horse. Participants rated on a 10-point scale how likely they thought each horse was to win a race individually, and then ranked all the horses, in exchange for £0.50.

A one-way repeated measures ANOVA showed that some descriptions were more likely to be rated as winning horses ($F(3, 116) = 10.64$, $p < .001$, $\hat{\eta}^2_G = .216$, 90% CI $[.104, .311]$). Therefore, we edited those passages and had a new sample of 30 British English speaker rate them. A one-way repeated measures ANOVA showed no differences in each horse's rated likelihood of winning a race ($F(3, 116) = 0.76$, $p = .516$, $\hat{\eta}^2_G = .019$, 90% CI

[.000, .057]), nor in the order in which they were ranked (all $|t| < 2$). The final set of descriptions for the experiment can be found in Table 1. Each description was paired with one of the four visual stimuli.

We recorded a British English native speaker (Darrington area) and a non-native (L1: Italian) English speaker to create the auditory stimuli. Both speakers were female and of similar age. Passages were recorded one at a time. To elicit naturally disfluent recordings, both speakers were instructed to read the passages silently and then were recorded as they tried to recall the passage from memory. To avoid differences between the descriptions provided by the speakers, they were allowed to look at the descriptions as they spoke if they could not remember the continuation. We edited these recordings using Audacity to ensure that the recordings of each speaker had similar numbers of disfluencies in similar locations, by cross-splicing different recordings (recordings can be found on OSF). To create the fluent counterpart of each description, filled and mid-utterance silent pauses were excised, and elongations and between-clause silent pauses were reduced using the 'Tempo' function. The final auditory experimental stimuli consisted, for each speaker, of two descriptions of each horse (one disfluent, one fluent), resulting in sixteen recordings. Description, fluency, speaker, and order of presentation were counterbalanced in a Latin Square design, resulting in 24 lists.

To ensure that the resulting descriptions were perceived to be natural (i.e., edits to the 'fluent' audios were not obvious) and that they were perceived as differing in fluency (i.e., disfluent and 'fluent' versions were distinguishable), we validated them in a sample of 48 British English participants, who did not take further part in the study. Participants were allocated to one of the 24 experimental lists, to ensure that our validation procedure was similar to how participants encountered stimuli in the actual experiment. Following Bosker et al.'s procedure, participants were asked to rate each audio's fluency on a scale from 1 to 9 (1: not fluent at all, 9: very fluent). We instructed participants to rate fluency by considering silent and filled pauses, speech rate, and repairs, and to ignore speakers' accents and the

Table 1

Original description of each horse that speakers were asked to memorise and reproduce.

Horse	Description
Fire Walker	Fire Walker is looking strong thanks to his come-from-behind success in the Acomb Stakes. The impression given in both runs is that Fire Walker should handle the demands of the extra furlong and Charlie Hills is looking forward to the test. The trainer said "He's done really well for a little break, his work's been good and I couldn't be more pleased with him".
Silver Sky	Silver Sky, a runner-up of a seven-furlong maiden at Naas on his debut last month, the son of Invincible Spirit ran crack French colt Persian King to a neck in the Group 3 Autumn Stakes over today's trip at Newmarket two weeks ago and his trainer believes he has done well since. O'Brien said "Silver Sky is a fine big colt and a talented one".
Apocalypse	Apocalypse has put in a string of consistent performances, most recently finishing third to Norway in the Zetland. "He's had a very solid year" said trainer Archie Watson. "He ran a good race in the Zetland, beaten only a length and a quarter, and I think the field here is of a similar level so I'm more than happy for him to take his chance".
Black Blade	Black Blade proved the market all wrong as the complexion of the 6.5-furlong novice race changed dramatically in the final two furlongs, with the Rebel Racing premier-owned newcomer under Tom Queally collaring long-time leaded Monsieur Noir. Spencer said: "He did it well. He's a nice horse. We always thought he had a bright future".

Table 2

Mean (standard deviation) ratings of native and non-native speakers' fluent and disfluent recordings for fluency, naturalness, and accentedness, on a 9-point scale where lower values indicate less fluent, less natural, and less accent respectively.

Speaker	Delivery	Fluency	Naturalness	Accentedness
Native Speaker	Fluent	7.62 (1.59)	5.44 (2.62)	4.38 (2.22)
	Disfluent	6.44 (2.36)	5.73 (2.57)	4.00 (1.97)
Non-native Speaker	Fluent	6.25 (1.73)	5.94 (1.83)	7.02 (1.59)
	Disfluent	5.21 (1.75)	5.83 (1.99)	7.17 (1.51)

content of their speech. Participants additionally rated on a 9-point scale each recording's naturalness (defined as how likely it was that the audio had been recorded in one go; 1: not unlikely at all; 9: very likely), and accentedness (while ignoring the perceived speaker's proficiency in the language; 1: not accented at all, 9: very accented). We additionally asked participants to guess the speakers' country of origin. At the end of the task, participants were further asked how often they interacted with native and non-native English speakers (on a 9-point scale, 1: never, 9: always) and were allowed to report if they noticed anything odd in the auditory stimuli.

Table 2 shows the means (and standard deviations) of participants' ratings of fluency, naturalness, and accentedness. A linear mixed model for fluency ratings with fixed effects of fluency (treatment-coded, reference: fluent), speaker's linguistic background (treatment-coded, reference: native speaker), and their interaction, with random intercepts by participant and by horse description, showed that fluency ratings differed significantly for the fluent and disfluent conditions ($\hat{\beta} = -1.19$, 95% CI $[-1.79, -0.58]$, $t = -3.84$). The non-native speaker was perceived as more disfluent than the native speaker ($\hat{\beta} = -1.37$, 95%

CI $[-1.97, -0.78]$, $t = -4.51$), in line with previous findings (Bosker, Quené, Sanders, & Jong, 2014; Pinget et al., 2014), but the interaction between the two variables was not significant ($\hat{\beta} = 0.15$, 95% CI $[-0.49, 0.78]$, $t = 0.45$). An identical model for naturalness ratings showed no significant differences by fluency ($\hat{\beta} = 0.29$, 95% CI $[-0.52, 1.11]$, $t = 0.70$), speaker's linguistic background ($\hat{\beta} = 0.50$, 95% CI $[-0.27, 1.27]$, $t = 1.27$) or their interaction ($\hat{\beta} = -0.40$, 95% CI $[-1.49, 0.70]$, $t = -0.71$). Finally, an identical model for accentedness only showed a significant difference for speaker's linguistic background ($\hat{\beta} = 2.65$, 95% CI $[2.04, 3.26]$, $t = 8.49$).

Procedure

Stimuli were presented using JsPsych (Leeuw et al., 2023), hosted on MindProbe (via JATOS, Lange et al., 2015). The task began with a cover story introducing two horse racing tipsters. Participants were told the tipsters would provide information about the four most popular horses competing in an upcoming race. The cover story explained that the two tipsters were well-known experts in the field, and added that one of the speakers was a non-native English speaker (without specifying the nationality of either speaker), introducing the element of the speakers' linguistic backgrounds as well as the factor of competence.

At the beginning of the experiment, participants were shown four pictures of the horses that they had been told would take part in the race, alongside their names. Participants were instructed to distribute one hundred pounds in betting money across the four horses based on the likelihood they thought each horse had of winning: They could split the bets as they wished, and they did not have to spend all the money. Each participant was randomly assigned to one out of 24 lists, so that they would listen to each speaker twice, one in each fluency level. The order in which horses were presented was randomised. In each trial, participants listened as one speaker described a given horse's performance. Once the playback stopped, participants were asked to place a bet by typing a number on a web form. Participants could only move to the next horse's description once they had placed a bet.

Participants were allowed to modify their previous bets every time they heard a new description. If the sum of bets made at any point was more than the allotted maximum, they were asked to re-distribute their bets until the total was below or equal to one hundred pounds.

After the betting round was complete, participants completed a questionnaire similar to that of Foucart et al. (2020) to measure their language attitudes towards the native and the non-native speaker. For each speaker, participants answered six questions measuring affect (three questions for negative affect and three questions for positive affect), five questions measuring solidarity, five questions measuring status, and one question each for comprehensibility, accentedness, fluency, and trustworthiness. Each question used a 9-point scale. Participants first answered questions about, at random, the native or non-native speaker, with the order of presentation of dimensions being randomised. They then answered the same questions for the remaining speaker. We also asked participants to guess the countries of origin of our native and non-native speakers, and from which speaker they would like to learn about horse races in the future. Additional questions included ratings on a 9-point scale of how natural the audio sounded (1: unnatural, edited; 9: natural, unedited), and two questions measuring participants' exposure to native and non-native accented English (1: never; 9: always). Likewise, we measured participants' previous experience with betting and their perceived knowledge of horse races (two questions: Whether they had bet on horse races in the past, and to rate on a 5-point scale, from 'Strongly disagree' to 'Strongly agree', how closely they identified with the statement 'I am an expert on horse races'). Finally, we included an open-ended question for participants to report what guided their decision-making, as well as their perception of the experiment's aim.

Results

All data pre-processing and analyses were carried out in R version 4.4.1 (R Core Team, 2024), using the packages *tidyverse* version 2.0.0 (Wickham et al., 2019), *ggplot* version

Table 3

Mean (standard deviation) of money bet by manner of delivery and speaker's linguistic background.

Speaker	Delivery	Money bet
Native	Fluent	24.24 (13.92)
	Disfluent	20.66 (12.23)
Non-native	Fluent	24.15 (13.61)
	Disfluent	22.31 (13.81)

3.5.1 (Wickham, 2016), and *wesanderson* version 0.3.7 for data wrangling and visualization. *lme4* version 1.1-35.5 (Bates et al., 2015) was used for data analysis, and *papaja* version 0.1.2 (Aust & Barth, 2023) for manuscript write-up. Scripts can be found at https://osf.io/zsut7/.

Pre-registered analyses

Betting behaviour

Table 3 shows the mean amount of money bet per speaker and fluency condition. On average, participants bet £24 on when information was provided fluently while disfluent instructions seem to lower the amount of money bet.

We modelled participants' betting behaviour using a linear mixed model. We modelled money bet on each horse, taken from the final amounts submitted in the experiment, after all four descriptions had been heard and valid responses (summing to at most 100 pounds) had been recorded. The model included fixed effects of fluency (sum-coded; fluent coded as -0.5, disfluent as +0.5), speaker's linguistic background (sum-coded, native coded as -0.5; non-native coded as +0.5), and their interaction. The maximal model (Barr et al., 2013), with random intercepts by-participant and by-item, with random slopes for fluency and speaker's linguistic background by-participant, and for fluency

by-item, failed to converge. We first dropped the random intercept by-participant, as most participants used all the money and thus there was no variance in their intercept. The final model only included a random intercept by-item.

Our model showed a main effect of fluency, whereby participants placed lower bets following disfluent descriptions compared to their fluent counterparts ($\hat{\beta} = -2.71$, 95% CI $[-4.04, -1.38]$, $t(1, 433) = -3.99$, $p < .001$). There was no main effect of speaker ($\hat{\beta} = 0.78$, 95% CI $[-0.55, 2.11]$, $t(1, 433) = 1.15$, $p = .251$) and no interaction between manner of delivery and speaker ($\hat{\beta} = 1.74$, 95% CI $[-0.92, 4.40]$, $t(1, 433) = 1.28$, $p = .200$)[1].

Language attitudes and post-experimental questionnaire

Table 4 depicts the means (and standard deviation) of ratings in each of the measures of interest by speaker's linguistic background. Constructs measured with more than one question (affect, status, solidarity) were obtained by calculating the average score. For the *Affect* dimension, we reverse-scored items measuring negative affect. Cronbach's alpha showed that the scores for these attitudes were reliable ($\alpha_{affect} = 0.80$, $\alpha_{status} = 0.93$, $\alpha_{solidarity} = 0.87$). We explored differences between evaluations of the native and the non-native speaker in these three social dimensions as well as on Comprehensibility, Accentedness, Fluency, and Trustworthiness via paired t-test using Bonferroni correction for p values.

Analyses showed that speakers were rated differently across all seven dimensions (see Table 4). The largest differences were, unsurprisingly, in the linguistic variables for comprehensibility, fluency and accentedness. The native speaker was rated as more comprehensible, more fluent and less accented. Following our pre-registered analysis, we

[1] An identical model including the 281 excluded participants showed a main effect of fluency ($\hat{\beta} = -2.92$, 95% CI $[-3.91, -1.92]$, $t(2, 557.08) = -5.75$, $p < .001$), and a main effect for speaker ($\hat{\beta} = 1.21$, 95% CI $[0.21, 2.20]$, $t(2, 557.09) = 2.38$, $p = .017$), but no significant interaction ($\hat{\beta} = 0.93$, 95% CI $[-1.06, 2.92]$, $t(2, 557.12) = 0.92$, $p = .358$)

Table 4

Average score (and standard deviation) in each dimension by speaker, and paired t-test for each dimension between speakers.

Dimension	Native Speaker	Non-native Speaker	p-value	d	t(359.00)
Comprehensibility	7.62 (1.57)	5.51 (2.14)	<.001	2.11	15.99
Accentedness	5.46 (2.1)	6.84 (1.37)	<.001	-1.39	-11.51
Fluency	7.97 (1.43)	6.11 (1.83)	<.001	1.87	17.15
Affect	5.77 (1.49)	6.05 (1.32)	= 0.003	-0.29	-3.54
Status	5.95 (1.52)	6.43 (1.31)	<.001	-0.48	-6.55
Solidarity	5.67 (1.48)	6.41 (1.24)	<.001	-0.74	-9.63
Trustworthiness	6.00 (1.61)	6.36 (1.41)	<.001	-0.36	-4.12

included these seven variables and their interaction with speaker identity to our previous model to explore whether speakers' evaluations could further explain participants' betting behaviours. This second model improved model fit ($\chi^2(14) = 33.67$, $p < .01$). The model showed a main effect of manner of delivery ($\hat{\beta} = -2.71$, 95% CI $[-4.03, -1.40]$, $t(1, 436.00) = -4.04$, $p < .001$). Regarding the linguistic variables, the higher the ratings on comprehensibility, the more money bet ($\hat{\beta} = 0.57$, 95% CI $[0.03, 1.11]$, $t(1, 436.43) = 2.08$, $p = .038$) and paralleling the effect of manner of delivery, higher ratings of fluency yielded higher bets ($\hat{\beta} = -0.61$, 95% CI $[-1.19, -0.02]$, $t(1, 436.08) = -2.04$, $p = .041$). As to the non-linguistic variables, affect was marginally significant ($\hat{\beta} = 0.67$, 95% CI $[-0.01, 1.36]$, $t(1, 436.06) = 1.92$, $p = .055$). The model also showed an interaction between status and speaker identity ($\hat{\beta} = -1.91$, 95% CI $[-3.57, -0.24]$, $t(1, 436.04) = -2.25$, $p = .025$): For the native speaker, higher ratings in status yielded higher money bet, while the opposite pattern was found for the non-native speaker.

Exploratory analysis

We additionally explored participants' preferences to learn from either speaker in the future. There was a slight preference to learn from the native speaker, with 206 participants choosing this speaker compared to 154 who preferred the non-native speaker. A χ^2 test of goodness of fit showed that this difference in preferences was reliable ($\chi^2(1) = 7.51$, $p < .01$).

Following previous research suggesting the role of exposure to foreign-accented speech in its role to trust and comprehensibility, we explored whether exposure to native and non-native speakers played a role in participants' betting behaviour beyond that of speaker identity and manner of delivery. A similar linear mixed model to that employed in our main analysis, but with the addition of the interaction between exposure and speaker, showed that it did not affect participants' betting behaviour ($\hat{\beta} = -0.35$, 95% CI $[-1.35, 0.66]$, $t(1, 431.10) = -0.68$, $p = .499$) nor was it a main effect ($\hat{\beta} = 0.42$, 95% CI $[-0.08, 0.92]$, $t(1, 431.25) = 1.63$, $p = .103$).

Discussion

Speech (dis)fluency has been previously shown to impact how confident in their knowledge a speaker is judged to be (Brennan & Williams, 1995). However, speakers can be disfluent for reasons other than lack of confidence: For example, speaking in one's second language is also associated with an increase in disfluencies (De Jong et al., 2015; Derwing et al., 2009). In the present experiment, we investigated whether speaker disfluency would have behavioral consequences where measurement of their knowledgeability was implicit, and whether speaker nativeness would be taken into account when listeners made their choices. Our findings suggest that manner of delivery, in the form of fluency, was the sole factor that guided participants' behavior: Participants placed lower bets on horses that were described disfluently, regardless of the accent of the speaker.

These findings reinforce the idea that listeners are sensitive to how an utterance is produced. In Brennan and Williams' (1995) study, participants explicity rated factual

responses which included hesitation phenomena as less likely to be correct, or in other words, made inferences about the degree of confidence the speaker had in their knowledge. In the present experiment, participants were selected to have little a priori knowledge that could have guided their betting behaviour. Moreover, the utterances they heard were semantically identical and, within speakers, acoustically matched apart from the excision of disfluencies. In the absence of other factors, a reasonable conclusion is that participants' betting decisions were predicated on inferences about the confidence with which each horse was described; and these inferences were affected by the speaker's disfluencies.

There was no evidence that participants' decisions were swayed by the fact that a non-native speaker might be expected to be disfluent for other reasons. This pattern differs from what we predicted given previous research suggesting that stereotypes about speaker's linguistic abilities have consequences for whether and how disfluencies affect speech comprehension (Arnold et al., 2007; e.g., Bosker, Quené, Sanders, & De Jong, 2014; Heller et al., 2015), and the fact that errors produced by non-native speakers are less likely to trigger otherwise negative interpretations (e.g., Fairchild et al., 2020; Fairchild & Papafragou, 2018; Ip & Papafragou, 2022; Lorenzoni et al., 2022) possibly because native listeners expect them to produce those (Lev-Ari, 2015). This is particularly remarkable given that in our post-experimental questionnaire, the native speaker was more likely to be chosen as someone participants would like to learn from about horse races in the future.

Another recent study failed to find an interaction between disfluencies and nativeness when judging speaker knowledge, although an interaction was evident in ratings of willingness to grant a request (Matzinger et al., 2023). Matzinger et al. attributed their findings to the degree with which a listener needs to model the speaker: Assessing a speaker's knowledge is not as socially engaging as other communicative contexts (e.g., granting a request) and consequently, inferring an interlocutor's cognitive state is not relevant. However, there are several differences between their paradigm and ours that cast

doubt on whether this explanation applies to our results. Specifically, in Matzinger et al. participants overheard a conversation, and were asked explicitly to rate their perceptions of each speaker's knowledge or confidence. In the present experiment, participants listened to speakers talking in isolation; most importantly, assessing speakers' knowledgeability was critical to task performance.

An aspect of the current experiment that might have affected the outcome is that both speakers were introduced to participants as knowledgeable tipsters. The fact that the speakers were purported to be authoritative figures may have been more salient than other features of their identities, including that of being a native or non-native speaker of English. Indeed, beliefs about a speaker's expertise have been shown to guide perceptions of their certainty (Mol et al., 2013). The preference to learn in future occasions from the native speaker might reflect ease of comprehension, rather than an implicit negative bias towards the non-native speaker, and particularly, a diminished perception of any inference concerning the non-native speaker's competence.

This experiment introduced a novel approach to measuring how knowledgeable the speaker is perceived to be. While previous studies had participants explicitly rate a speaker on different dimensions (e.g., knowledgeability, trust) in non-social contexts, the present study avoids explicit judgement and, instead, relies on behaviors which are highly likely to index an implicit judgement. However, a potential shortcoming of our design has to do with how people approached the betting system: Most participants followed a conservative approach and distributed their bets fairly evenly. It is, therefore, possible that in situations of less uncertainty (e.g., a scenario where speakers are deemed to have more knowledge) or where participants are rewarded only for allocating money to the winning horse, speaker perception could yield a larger effect on participants' behaviours.

Overall, our study shows that disfluency is a cue that listeners take as an index of a speaker's (lack of) knowledgeability. Importantly, in situations of uncertainty where

individuals may lack information to optimise their behavior, the presence of disfluency biases their otherwise rational behavior. This highlights that individuals are attuned to the different ways in which a speaker's mental state is signalled in speech and demonstrates how non-verbal aspects of spoken messages can directly affect people's subsequent behaviours.

References

Armstrong, S. L., Gleitman, L. R., & Gleitman, H. (1983). What some concepts might not be. *Cognition*, *13*(3), 263–308.

Arnold, J. E., Kam, C. L. H., & Tanenhaus, M. K. (2007). If you say thee uh you are describing something hard: The on-line attribution of disfluency during reference comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(5), 914–930. https://doi.org/10.1037/0278-7393.33.5.914

Aust, F., & Barth, M. (2023). *papaja: Prepare reproducible APA journal articles with R Markdown*. https://github.com/crsh/papaja

Barlow, S., Beardsley, G., Bsharah, Z., Crofts, R., De La Rosa, C., Gutierrez, A., Highfill, C., Johnson, A. G. W., Johnson, C., Johnson, J., & al., et. (2024). The effects of exposure and explicit stereotypes on veracity judgments of Polish-accented English speech: A preregistered close replication and extension of Boduch-Grabka & Lev-Ari (2021). *Studies in Second Language Acquisition*, 1–17. https://doi.org/10.1017/S0272263124000123

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Barr, D. J., & Seyfeddinipur, M. (2010). The role of fillers in listener attributions for speaker disfluency. *Language and Cognitive Processes*, *25*(4), 441–455. https://doi.org/10.1080/01690960903047122

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bazzi, L., Brouwer, S., & Foucart, A. (2022). The impact of foreign accent on irony and its consequences on social interaction. *Journal of Multilingual and Multicultural Development*, 1–13. https://doi.org/10.1080/01434632.2022.2126487

Bergmann, C., Sprenger, S. A., & Schmid, M. S. (2015). The impact of language

co-activation on L1 and L2 speech fluency. *Acta Psychologica*, *161*, 25–35. https://doi.org/10.1016/j.actpsy.2015.07.015

Boduch-Grabka, K., & Lev-Ari, S. (2021). Exposing Individuals to Foreign Accent Increases their Trust in What Nonnative Speakers Say. *Cognitive Science*, *45*(11), e13064. https://doi.org/10.1111/cogs.13064

Bosker, H. R., Quené, H., Sanders, T., & De Jong, N. H. (2014). Native 'um's elicit prediction of low-frequency referents, but non-native 'um's do not. *Journal of Memory and Language*, *75*, 104–116. https://doi.org/10.1016/j.jml.2014.05.004

Bosker, H. R., Quené, H., Sanders, T., & Jong, N. H. de. (2014). The perception of fluency in native and nonnative speech. *Lang. Learn.*, *64*(3), 579–614. https://doi.org/10.1111/lang.12067

Brennan, S. E., & Williams, M. (1995). The Feeling of Another's Knowing: Prosody and Filled Pauses as Cues to Listeners about the Metacognitive States of Speakers. *Journal of Memory and Language*, *34*(3), 383–398. https://doi.org/10.1177/0261927X10387100

Butterworth, H. (2019). *How much do they, um, know?: Disfluencies in speech as cues to speakers' knowledge.* University of Edinburgh.

Caffarra, S., Michell, E., & Martin, C. D. (2018). The impact of foreign accent on irony interpretation. *PLoS One*, *13*(8), e0200939. https://doi.org/10.1371/journal.pone.0200939

Christenfeld, N. (1995). Does it hurt to say um? *Journal of Nonverbal Behavior*, *19*, 171–186. https://doi.org/10.1007/BF02175503

De Jong, N. H., Groenhout, R., Schoonen, R., & Hulstijn, J. H. (2015). Second language fluency: Speaking style or proficiency? Correcting measures of second language fluency for first language behavior. *Applied Psycholinguistics*, *36*(2), 223–243. https://doi.org/10.1017/S0142716413000210

DeBruine, L. M., & Barr, D. J. (2021). Understanding mixed-effects models through data simulation. *Advances in Methods and Practices in Psychological Science*, *4*(1),

2515245920965119. https://doi.org/10.1177/251524592096511

Derwing, T. M., Munro, M. J., Thomson, R. I., & Rossiter, M. J. (2009). The relationship between L1 fluency and L2 fluency development. *Studies in Second Language Acquisition*, *31*(4), 533–557. https://doi.org/10.1017/S0272263109990015

Dragojevic, M., & Giles, H. (2016). I Don't like You because You're Hard to Understand: The Role of Processing Fluency in the Language Attitudes Process. *Human Communication Research*, *42*(3), 396–420. https://doi.org/10.1111/hcre.12079

Fairchild, S., Mathis, A., & Papafragou, A. (2020). Pragmatics and social meaning: Understanding under-informativeness in native and non-native speakers. *Cognition*, *200*, 104171. https://doi.org/10.1016/j.cognition.2019.104171

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

Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D. M., & Perrett, D. I. (2005). Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Animal Behaviour*, *69*(3), 561–568. https://doi.org/10.1016/j.anbehav.2004.06.012

Foucart, A., Costa, A., Morís-Fernández, L., & Hartsuiker, R. J. (2020). Foreignness or processing fluency? On understanding the negative bias toward foreign-accented speakers. *Language Learning*, *70*(4), 974–1016. https://doi.org/10.1111/lang.12413

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

Gkalitsiou, Z., & Werle, D. (2023). Speech disfluencies in bilingual Greek-English young adults. *Journal of Fluency Disorders*, *78*, 106001. https://doi.org/10.1016/j.jfludis.2023.106001

Gluszek, A., & Dovidio, J. F. (2010). The Way They Speak: A Social Psychological

Perspective on the Stigma of Nonnative Accents in Communication. *Personality and Social Psychology Review*, *14*(2), 214–237. https://doi.org/10.1177/108886830935928

Gluszek, A., Newheiser, A.-K., & Dovidio, J. F. (2011). Social Psychological Orientations and Accent Strength. *Journal of Language and Social Psychology*, *30*(1), 28–45. https://doi.org/10.1177/0261927X10387100

Goupil, L., Ponsot, E., Richardson, D., Reyes, G., & Aucouturier, J.-J. (2021). Listeners' perceptions of the certainty and honesty of a speaker are associated with a common prosodic signature. *Nature Communications*, *12*(1), 861. https://doi.org/10.1038/s41467-020-20649-4

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, *109*(1), 3. https://doi.org/10.1037/0033-295X.109.1.3

Grosjean, F., & Deschamps, A. (1975). Analyse contrastive des variables temporelles de l'anglais et du français: Vitesse de parole et variables composantes, phénoménes d'hésitation. *Phonetica*, *31*(3-4), 144–184. https://doi.org/10.1159/000259667

Guyer, J. J., Fabrigar, L. R., & Vaughan-Johnston, T. I. (2019). Speech Rate, Intonation, and Pitch: Investigating the Bias and Cue Effects of Vocal Confidence on Persuasion. *Personality and Social Psychology Bulletin*, *45*(3), 389–405. https://doi.org/10.1177/0146167218787805

Hanuíková, A., Alphen, P. M. van, Goch, M. M. van, & Weber, A. (2012). When One Person's Mistake Is Another's Standard Usage: The Effect of Foreign Accent on Syntactic Processing. *Journal of Cognitive Neuroscience*, *24*(4), 878–887. https://doi.org/10.1162/jocn_a_00103

Heller, D., Arnold, J. E., Klein, N., & Tanenhaus, M. K. (2015). Inferring Difficulty: Flexibility in the Real-time Processing of Disfluency. *Language and Speech*, *58*(2), 190–203. https://doi.org/10.1177/0023830914528107

Ip, M. H. K., & Papafragou, A. (2022). Integrating Non-Native Speaker Identity in Semantic

and Pragmatic Processing. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *44*. https://doi.org/escholarship.org/uc/item/1qk4z377

Jiang, X., & Pell, M. D. (2015). On how the brain decodes vocal cues about speaker confidence. *Cortex*, *66*, 9–34. https://doi.org/10.1016/j.cortex.2015.02.002

Jiang, X., & Pell, M. D. (2016a). Neural responses towards a speaker's feeling of (un) knowing. *Neuropsychologia*, *81*, 79–93. https://doi.org/10.1016/j.neuropsychologia.2015.12.008

Jiang, X., & Pell, M. D. (2016b). The feeling of another's knowing: How "mixed messages" in speech are reconciled. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(9), 1412. https://doi.org/10.1037/xhp0000240

Jiang, X., & Pell, M. D. (2017). The sound of confidence and doubt. *Speech Communication*, *88*, 106–126. https://doi.org/10.1016/j.specom.2017.01.011

Lange, K., Kühn, S., & Filevich, E. (2015). "Just another tool for online studies" (JATOS): An easy solution for setup and management of web servers supporting online studies. *PloS One*, *10*(6), e0130834. https://doi.org/10.1371/journal.pone.0130834

Leeuw, J. R. de, Gilbert, R. A., & Luchterhandt, B. (2023). jsPsych: Enabling an open-source collaborative ecosystem of behavioral experiments. *Journal of Open Source Software*, *8*(85), 5351. https://doi.org/10.21105/joss.05351.

Lev-Ari, S. (2015). Comprehending non-native speakers: Theory and evidence for adjustment in manner of processing. *Frontiers in Psychology*, *5*, 1546. https://doi.org/10.3389/fpsyg.2014.01546

Lev-Ari, S., & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology*, *46*(6), 1093–1096. https://doi.org/10.1016/j.jesp.2010.05.025

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

598 Matzinger, T., Pleyer, M., & Żywiczyński, P. (2023). Pause Length and Differences in
599 Cognitive State Attribution in Native and Non-Native Speakers. *Languages*, *8*(1), 26.
600 https://doi.org/10.3390/languages8010026
601 McConnell, A. R., & Leibold, J. M. (2001). Relations among the Implicit Association Test,
602 Discriminatory Behavior, and Explicit Measures of Racial Attitudes. *Journal of*
603 *Experimental Social Psychology*, *37*(5), 435–442. https://doi.org/10.1006/jesp.2000.1470
604 Mol, L., Kuhlen, A., Van der Steen, R., & Obbens, M. (2013). Beliefs about a speaker affect
605 feeling of another's knowing. *Proceedings of the Annual Meeting of the Cognitive Science*
606 *Society*, *35*. https://doi.org/escholarship.org/uc/item/8w03w4ch
607 Norton-Ford, J. D., & Hogan, D. R. (1980). Role of Nonverbal Behaviors in Social
608 Judgments of Peers' Assertiveness. *Psychological Reports*, *46*(3_suppl), 1085–1086.
609 https://doi.org/10.2466/pr0.1980.46.3c.1085
610 Payne, B. K., & Hannay, J. W. (2021). Implicit bias reflects systemic racism. *Trends in*
611 *Cognitive Sciences*, *25*(11), 927–936. https://doi.org/10.1016/j.tics.2021.08.001
612 Pinget, A. F., Bosker, H. R., Quené, H., & De Jong, N. H. (2014). Native speakers'
613 perceptions of fluency and accent in L2 speech. *Language Testing*, *31*(3), 349–365.
614 https://doi.org/10.1177/026553221452
615 Pittam, J., & Gallois, C. (1986). Predicting Impressions of Speakers from Voice Quality:
616 Acoustic and Perceptual Measures. *Journal of Language and Social Psychology*, *5*(4),
617 233–247. https://doi.org/10.1177/0261927X8600500401
618 R Core Team. (2024). *R: A language and environment for statistical computing*. R
619 Foundation for Statistical Computing. https://www.R-project.org/
620 Rakić, T., Steffens, M. C., & Mummendey, A. (2011). When it matters how you pronounce
621 it: The influence of regional accents on job interview outcome. *British Journal of*
622 *Psychology*, *102*(4), 868–883. https://doi.org/10.1111/j.2044-8295.2011.02051.x
623 Shriberg, E. E., Bates, R., & Stolcke, A. (1997). A prosody only decision-tree model for
624 disfluency detection. *Fifth European Conference on Speech Communication and*

Technology. https://doi.org/10.21437/eurospeech.1997-626

Shriberg, E. E., & Lickley, R. J. (1993). Intonation of Clause-Internal Filled Pauses. *Phonetica*, *50*(3), 172–179. https://doi.org/10.1159/000261937

Smith, V. L., & Clark, H. H. (1993). On the Course of Answering Questions. *Journal of Memory and Language*, *32*(1), 25–38. https://doi.org/10.1006/jmla.1993.1002

Souza, A. L., & Markman, A. B. (2013). Foreign accent does not influence cognitive judgments. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *35*. https://doi.org/escholarship.org/uc/item/8hd9v4ff

Stocker, L. (2017). The Impact of Foreign Accent on Credibility: An Analysis of Cognitive Statement Ratings in a Swiss Context. *Journal of Psycholinguistic Research*, *46*, 617–628. https://doi.org/10.1007/s10936-016-9455-x

Uleman, J. S., Adil Saribay, S., & Gonzalez, C. M. (2008). Spontaneous Inferences, Implicit Impressions, and Implicit Theories. *Annu. Rev. Psychol.*, *59*, 329–360. https://doi.org/10.1146/annurev.psych.59.103006.093707

Wetzel, M., Zufferey, S., & Gygax, P. (2021). Do non-native and unfamiliar accents sound less credible? An examination of the processing fluency hypothesis. *Journal of Articles in Support of the Null Hypothesis*, *17*(2). https://doi.org/10.48350/152025

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., ... Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686