In situations where the facts are not known, we usually need to decide whether to believe what we are being told by assessing either the content of the communicated information or the source of the information. Individuals are equipped with a set of cognitive abilities to monitor whether they are being misinformed (i.e., epistemic vigilance), which has been argued to be crucial in the development of culture and language (Sperber et al., 2010). The trustworthiness of a speaker can be thought to be assessed in two dimensions: How confident (or certain) a speaker is in the knowledge they have and how willing they are to share honest information. Although this online assessment may seem costly, there is considerable evidence suggesting that it might not be so: Even young children can assess the reliance of a source, for example (Poulin-Dubois & Brosseau-Liard, 2016). However, assuming a speaker’s honesty, the attribution of certainty can be biased by several, and potentially irrelevant, factors ranging from the situational (particularities about the speaker) to the ephemeral (particularities about the details of speech).

There is a growing body of literature showing that certain speech features impact listeners’ perceptions of a speaker’s certainty. Voices with a slow speech rate, low amplitude, and large F0 range are less likely to be rated as confident, and are associated with distinct neural responses in the listener (Jiang & Pell, 2015, 2016a, 2016b). Similarly, distinct prosodic contours can affect 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). Within disfluencies, filled pauses (e.g., ‘uh’ or ‘um’) have been shown to be used as indices to evaluate a speaker’s trustworthiness, both in terms of confidence (Brennan & Williams, 1995) and of honesty (Arciulli et al., 2011; Loy et al., 2017).

Brennan and Williams (1995) demonstrated that silent and filled (e.g., uh 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), a metric of perceived speaker’s certainty (or confidence). 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, that is, listeners interpreted that the speaker was sure they did not know the answer. Brennan and Williams (1995) 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 of a question (see also Smith & Clark, 1993 and Kirkland et al., 2023, for similar findings). In fact, listeners are attentive to speakers’ displays in both the visual and auditory modalities (Swerts & Kramer, 2005), and this sensitivity to cues associated with a speaker’s certainty seems to arise from an early age (Hübscher et al., 2017; Kramer & Swerts, 2005; Mori & Pell, 2017).

One explanation for these findings may have its roots in epistemic vigilance. The set of cues thus far discussed are likely to arise in situations where the speaker in under cognitive effort. For disfluencies, their production is increased when speakers experience cognitive load (Bortfeld et al., 2001) such as when describing a topic they are less familiar with (Merlo & Mansur, 2004), when having difficulties in word retrieval (Hartsuiker & Notebaert, 2010; Pistono & Hartsuiker, 2021; Schnadt & Corley, 2006). As interlocutors monitor each other’s mental states, when confronted with these cues, listeners may assume that the speaker is experiencing cognitive load and try to find the cause for such trouble. Depending on the conversational context, this may lead to the perception that the speaker is not confident in what they are saying, and thus the information is less reliable (compared to when it is produced fluently), working as a protective mechanism against misinformation.

A consequence of this hypothesis is that listeners may be more lenient with a speaker when they are disfluent if there are alternative explanations for such disfluency. While, for example, a disfluency may be interpreted as a sign of deception in one speaker, it may be taken as a display of inattentiveness in another (provided that there are additional cues that support this model, e.g., King et al., 2018). In the case of disfluency, certain characteristics of a speaker may make them particularly prone to produce them: For example, speakers producing speech in their second language are generally more disfluency (Bergmann et al., 2015; Gkalitsiou & Werle, 2021) and are perceived as such by native listeners (Pinget et al., 2014), and therefore, a non-native speaker’s disfluencies may lead to a different interpretation than its native counterpart.

While there is a wealth of research suggesting that a speaker’s non-nativeness can trigger negative evaluations due to difficulties in processing speech with an accent different from one’s own (e.g., Dragojevic & Giles, 2016; Gluszek & Dovidio, 2010; Lev-Ari & Keysar, 2010 but cf. Souza & Markan, 2013; Stocker, 2017; Foucart, Costa, Morís-Fernández, & Hartsuiker, 2020; Foucart & Hartsuiker, 2021; Wetzel, Zufferey, & Gygax, 2021; Barlow et al., 2024), listeners may hold a model of non-native speakers that may facilitate this lenience hypothesis. For example, a non-native speaker’s failure to be informative is more likely to be taken as a sign of inability to produce the necessary information than as unwillingness to provide the information, even when participants were not told that the non-native speaker could struggle with language production (Fairchild et al., 2020). Importantly, a non-native underinformative speaker is preferred over a native one as an option to learn in the future (Fairchild et al., 2020; for similar findings see Fairchild & Papafragou, 2018; Lorenzoni et al., 2022; Ip & Papafragou, 2022).

The pattern of results described above might be attributable to expectations about non-native speakers. Specifically, individuals may hold a stereotype about non-native speakers whereby they are expected to be less linguistically competent than a native speaker (Fairchild & Papafragou, 2018; Lev-Ari, 2015). These stereotypes in turn drive how their speech is comprehended and interpreted. For example, when encountering an ambiguity in the speech signal (e.g., /fεri/, somewhere between ferry and fairy), listeners would rely more on the discourse context than on bottom-up signal to disambiguate the intended referent (Lev-Ari, 2015). 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 (Hanulíková et al., 2012).

This expectation of reduced linguistic ability may also translate onto how disfluencies affect speech comprehension. In an eye-tracking experiment, Bosker et al. (2014) found that listeners displayed anticipatory eye-movements towards objects with low-frequency labels (compared to high-frequency labels) if they were following instructions provided by a native speaker, allegedly because the production of such labels is cognitive costly and thus a likely explanation for the presence of a disfluency. In contrast, listeners showed no predictive behaviour if the speaker was producing speech in their second language. 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). This pattern of results suggests that not only listeners are sensitive to those cues that signal effort on the speaker’s mind, but they can also factor in different causes for such effort.

There is, however, little evidence exploring whether and how non-native speakers’ disfluencies trigger different evaluations in terms of confidence and thus trustworthiness compared to natives’. 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 (see also Houghton et al., 2024 for a similar pattern). It is worth noting that Matzinger et al. (2023) and Houghton et al. (2024) worked with silent pauses. As discussed by Houghton et al. (2024), silent and filled pauses serve different functions (e.g., Rose, 2019) and consequently, may yield different interpretations.

As is the case for Matzinger et al. (2023), most of studies evaluating the effects of manner of speech in the manner of fluency on perceptions of the speaker have used explicit ratings – for example, to rate particular traits of the speaker 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 correlate with how participants would behave (Greenwald, 1990; Greenwald et al., 2002). While experiments eliciting metalinguistic judgements (e.g., ‘How confident does this speaker sound?’) shed light on what factors may affect how a speaker is perceived, the lack of interaction on participants’ behalf entails that these findings may not predict how listeners would act. In other words: while Brennan and Williams (1995) demonstrated that people attribute less confidence to disfluent speakers, would that mean that participants would be less likely to follow their advice? This question is particularly relevant to disentangle whether and how perceptions of a speaker guide actual decision-making.

In contrast to explicit measures, implicit measures overcome these obstacles while still being able to predict behaviours. Examples of these measures are becoming more prevalent in the literature, with participants being asked to choose a speaker to learn from (e.g., Ip & Papafragou, 2022) or via economic games (Caballero & Pell, 2020; Torre et al., 2023). These paradigms allow researchers to explore how perceptions of the speaker impact listeners’ behaviour. To the authors best of knowledge, there is no prior study exploring how reduced perceived confidence (and thus, lower trust) impacts individuals’ behavours.

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 confident in their knowledge the speaker is). Instead, we take participants’ allocations of ‘betting money’ as an indirect measurement of their perceptions of the speaker and the consequent trust they put in their advice. 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 speaker’s confidence 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. As listeners monitor the speaker’s mental state in uncertain situations, we expect disfluency to be taken as reduced certainty and thereby decrease the money allocated to horses described disfluency. If listeners are sensitivity to the different causes for a speaker to be disfluent, a non-native speaker’s disfluency should not decrease money bet to the same extent that a native speaker’s. 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).

Discussion

Although this mechanism entails that listeners may misattribute the cause of the disfluency - for example, while speakers are no more disfluent when they produce lies (Loy et al., 2018; De keersmaecker et al., 2024), listeners consistently take filled pauses as indices of deception (e.g., Loy et al., 2017; King et al., 2018; Li et al., 2022).