Do non-native and unfamiliar accents sound less credible? An examination of the processing fluency hypothesis

Mathis Wetzel
Sandrine Zufferey
University of Bern

Pascal Gygax
University of Fribourg

Many studies have demonstrated that stimuli that are easy to process are generally better evaluated compared to stimuli that are harder to process. It is, however, an open question whether people speaking with a foreign accent are judged to be less truthful compared to native speakers due to the greater difficulty of decoding their speech. In this paper, we provide new data to this debate by comparing the credibility of speakers of French, both with a familiar or unfamiliar native accent, and with a familiar and unfamiliar foreign accent. Our results indicate that native Native-speakers do not evaluate statements uttered with a foreign-accent as less truthful compared to a native one.

University of Bern. Institute of French Language and Literature, Länggassstrasse 49, 3012 Bern, Switzerland
University of Fribourg. Departement of Psychology, Rue P.-A.-de-Faucigny 2, 1701 Fribourg, Switzerland

Corresponding author:
Mathis Wetzel, mathis.wetzel@rom.unibe.ch

Introduction

In everyday life, people constantly have to evaluate new information and unfamiliar situations in order to be able to make decisions. According to Oppenheimer (2008), the mechanism of evaluation and decision-making is, amongst other things, strongly influenced by the mode of presentation of the information that has to be evaluated. More specifically, when people are presented with a linguistic message that is easy to process and understand, they tend to form an overall more positive impression about its content than when the message is presented in a form that makes it hard to decode and interpret (Winkielman & Cacioppo, 2001). As a consequence, the subjective ease or difficulty with which people decode and interpret a stimulus has an impact on their propensity to accept or reject it.

When it comes to language use, one factor that typically increases the difficulty to process and interpret a message is the fact that it is pronounced with a foreign or a non-standard accent. As a result, it can be expected that speakers who have either a foreign or a non-standard accent may be considered to be less trustworthy compared to speakers who speak with a standard accent. While some studies in the literature have indeed found such an effect (e.g., Lev-Ari & Keysar, 2010; Dragojevic & Giles, 2016), others have failed to find it (e.g., Souza & Markman, 2013; Stocker, 2017). We argue that one crucial factor that may explain the difference between these studies, and that hasn't been considered so far, is the level of familiarity that speakers have with the foreign accent. Indeed, a foreign accent heard on a regular basis may not trigger the same processing difficulties compared to an accent that is highly unfamiliar.

In this paper, we assess the role of accent familiarity for people's propensity to trust or distrust a speaker, by including both an unfamiliar native accent to compare it to a familiar one, as well as two foreign accents, one highly familiar and another one highly unfamiliar. For this, we compared the perceived trustworthiness, for French-speaking people living in Switzerland, of four different types of speakers: (1) native speakers of French who have the same local Swiss French accent as the participants; (2) native-speakers of French with an unfamiliar accent from Québec; (3) non-native speakers who have a foreign but highly familiar German accent (Switzerland has a large population of German speakers, and there are important areas where both language speakers cohabit); (4) non-native speakers with a foreign and unfamiliar Finnish accent. Thanks to our comparison of both native and non-native, familiar and unfamiliar accents, our study enables us to disentangle the role of processing ease from the negative associations directed specifically towards to non-native speakers, therefore provide new data on an unresolved controversy in the literature.

The paper is structured as follows. We first present studies that have addressed the role of accents for speakers' credibility. We start by introducing the ease of processing effect and its influence on the ratings of various types of stimuli, before reviewing more specifically studies that have addressed its impact on the credibility of speakers with a foreign accent. We claim that the issue is far from settled, as some studies have evidenced such an effect while others failed to do so. Thereafter, we argue that speakers' familiarity with the non-native accent is potentially a relevant factor that has never been tested so far, and that could explain the differences observed between studies. In order to assess the role of familiarity, we present a new empirical study assessing various types native and non-native French speakers. We conclude that foreign accents, be they familiar or unfamiliar, do not seem to influence speakers' perception of trustworthiness. To substantiate our claim, we statistically demonstrate – calculating Bayes factor – that our data were sensitive enough to provide strong evidence for the null hypothesis. We finally discuss the impact of our findings and present future avenues of enquiry for studies on foreign accents and their impact for communication.

The role of accents for speaker credibility

The ease of processing effect

Many studies have demonstrated the existence of several factors that can facilitate or disrupt the fluency of mental processing. For example, at a visual level, when perceiving stimuli on a computer, increased color contrast or a clearly legible font facilitates the fluency of mental processing (e.g., Reber, Winkielman, & Schwarz, 1998; Novemsky, Dhar, Schwarz, & Simonson, 2007; Alter & Oppenheimer, 2007). Generally, items that are easier to process elicit more positive reactions (e.g., Alter & Oppenheimer, 2009) and therefore strongly but subconsciously influence the evaluations and the decisions that people make in everyday life. For example, when people are presented with names for food additives that are difficult to pronounce, they find them to be potentially more harmful compared to additives with names that are easy to pronounce (Song & Schwarz, 1999). In other studies, participants were found to prefer stimuli that were presented longer compared to shorter presentation, and also prefer stimuli that were presented twice over new stimuli, indicating that familiarity leads to subjective preferences (e.g., Reber et al., 1998; Winkielman & Cacioppo, 2001). The impact of processing ease is also explicitly reported in several studies to have an effect on decision making. For example, in the context of the stock market, Alter and Oppenheimer (2006) report that investors predicted a better short-term performance for (invented) companies that had ticker codes that were easy to pronounce compared (e.g., KAR) to ones that had abbreviations that were hard to pronounce (e.g., RDO).

While many studies have consistently demonstrated a wide range of daily-life implications linked to the ease of processing effect (see Alter & Oppenheimer, 2009 for an overview), there is an ongoing debate in linguistics about whether the higher mental effort involved in processing stimuli from speakers with a non-native accent influences their credibility in the mind of native speakers, as there is evidence in the literature that perceptual fluency has a great impact of the perceived truthfulness of an information. Reber and Schwarz (1999; as well as Unkelbach,

2007) reported for example that statements that were harder to read, due to manipulation of the font color in which the statements were presented, were believed to be less credible than statements that were presented in an easy-to-read color. It can be further assumed that the processing fluency effect can be triggered also by purely linguistics factors, as studies demonstrated for example that a message worded in simple syntax led to a better recall and increased its persuasiveness (Lowrey, 1998). In other experiments, items with simpler spelling felt psychologically nearer (Alter & Oppenheimer, 2007) and items with a simpler lexis led participants to believe that the author was more intelligent (Oppenheimer, 2006). However, studies that have assessed whether a non-standard speaker is perceived to be less credible due to the mental effort that their foreign accent is eliciting have provided mixed findings, as we now outline.

The credibility of speakers with a foreign accent

Lev-Ari and Keysar (2010) conducted one of the first studies assessing the role of foreign accent for speakers' perceived trustworthiness. In their experiment, 28 participants were asked to rate the truthfulness of 60 recordings of trivia statements that were presented in three conditions: (1) by native speakers; (2) by speakers with a mild foreign accent; (3) by speakers with a strong foreign accent. Participants were asked to judge the credibility of each statement by placing a cross on a continuous scale (of 14 cm) ranging from "definitely false" to "definitely true". In order to eliminate a possible bias due to pre-existing stereotypes about the trustworthiness of non-native speakers, Lev-Ari and Keysar (2010) told their participants that the speakers uttering the statements were only messengers but not the source of their content. Results from this experiment showed that the statements that were presented with the strong foreign accent were believed to be less truthful than ones that were presented by speakers that were either native speakers or had a mild foreign accent. In order to explain the difference between foreign speakers with a mild and a strong accent, Lev-Ari and Keysar (2010) proposed the *processing fluency theory*, according to which the high mental effort necessary to process a statement that is presented in a strong foreign accent elicits a negative impression that ultimately leads to a reduced credibility of the statement. This hypothesis found further support in a later study by Dragojevic and Giles (2016). In this study, participants were asked to evaluate utterances that were presented either in a standard variety (American English)[1] or in a non-standard variety (Panjabi English) by rating the speaker on a scale to 0 (*not at all*) to

7 (*very*) in relation to a set of adjectives that were grouped into two dimensions: status (e.g. intelligent, educated, smart) and solidarity (e.g. nice, pleasant, honest). In addition, processing fluency was disturbed by adding extra noise to some of the recordings for both standard and non-standard utterances. In a first experiment, participants evaluated the Punjabi accent with added noise more negatively in both dimensions compared to recordings that had no added noise. In contrast, for the standard variety, there was no difference between the judgements for recordings that had or had no added noise. In a second experiment, an additional step was taken to ensure that stereotypes did not bias the results. The standard and non-standard utterances were explicitly linked to specific origins (California and India) and the participants were asked to perform a memory task prior to rating the speaker. In this second experiment, participants attributed again lower status to the recordings with the Punjabi accent that had the added noise compared to recordings of the same accent with no added noise. This time however, recordings in the standard variety showed the same pattern, i.e. recordings in American English with added noise received a lower status judgment compared to the ones with no added noise. Throughout the two experiments, the recordings of the non-standard variety were rated lower on the dimension of solidarity than the standard variety. As predicted by the processing fluency theory, Dragojevic and Giles (2016) concluded that the added noise in the recordings reduced the ease of processing of the statements and consequently led to a more negative attitude towards these recordings. Additional support to this interpretation was provided in a follow-up study by Dragojevic, Giles, Beck, and Tatum (2017), who asked native speakers in two experiments to evaluate recordings of mild-accented and heavy-accented speakers. As in the previous study of Dragojevic and Giles (2016), the participants were asked to rate the speaker on a scale to 0 (*not at all*) to 7 (*very*) on a list of adjectives that were again grouped into the general dimensions of status and solidarity. While a heavy accent did not affect evaluations on the dimension of solidarity, the researchers reported that heavy-accented speakers were attributed again less status than mild-accented speakers.

In addition to that, Hanzlíková and Skarnitzl (2017) reproduced the initial experiment by Lev-Ari and Keysar (2010), but this time with non-native participants. Their results showed a positive bias towards native speakers, thus adding further support to the *processing fluency hypothesis*. Other studies supporting Lev-Ari and Keysar's findings were conducted by Leach and Da Silva (2012) and Evans and Michael (2014). In the former study, participants were asked to detect lies by rating utterances of native and non-native speakers in video recordings from a previous psychological experiment. The results showed a positive bias towards native speakers, indicating that the participants tended to believe that native speakers were more likely to tell the truth. In the latter study, participants were asked to judge the truthfulness of video recordings of native and non-native speakers. Results also showed that participants were negatively biased against non-native speakers. Although these two studies do support the processing fluency hypothesis, they did not, however, control for the potential bias of stereotypes.

While all studies presented so far lend support to the *processing*

---

[1] Although we mention *standard variety*, we acknowledge that 'American English' does in fact include many varieties, and thus many accents. Still, we consider the notion of standard variety, in this paper, in light of our respective samples. For example, Canadian French represents a non-standard variety since our participants were located in Fribourg, Switzerland. Conversely, had we tested in Québec, Canadian French would have been considered the *standard variety* (in line with Dragojevic & Giles, 2016, p. 398).

fluency hypothesis, several other studies have failed to replicate this effect. For example, based on the *processing fluency hypothesis*, Souza and Markman (2013) assessed whether added noise would change the credibility ratings of recorded statements provided by a native speaker (Experiment 1). Results showed that the added noise had no significant effect to the credibility ratings. In addition to this lack of effect, in a direct attempt to replicate Lev-Ari and Keysar's (2010) findings, Souza and Markman (2013) asked participants to rate the truthfulness of statements presented by a native and a non-native speaker (Experiment 2). Although the same methodology as in Lev-Ari and Keysar (2010) was used – only this time using a scale from 0 to 10 – participants did not judge the statements of native speakers as more credible that those of non-native ones. Souza and Markman (2013) concluded that the results of Lev-Ari and Keysar (2010) could have been caused by a type 1 error, in other words they could be a case of false positive. In another set of experiments, Stocker (2017) also tried to replicate Lev-Ari and Keysar's methodology with a higher number of participants and in two languages, French ($n = 194$) and German ($n = 184$), yet no credibility effect was found. Similarly, De Meo, Vitale, Pettorino, and Martin (2011) also failed to reproduce the difference in credibility between native and non-native accents, even though they increased the number of participants to 300. Similar to the methodology of Lev-Ari and Keysar, the De Meo et al. (2011) asked participants to evaluate the truthfulness of statements that were presented in a native and a non-native condition but found no direct interaction between a non-native accent and a perceived reduced credibility. However, De Meo et al. did observe some effect of non-standard suprasegmental acoustic features on credibility (see also De Meo, 2012).

Podlipský, Šimáčková, and Petráž (2016) compared the ratings provided by non-native and native-speaking participants. In this experiment, participants were asked to judge the credibility of statements that were presented by native and non-native speakers on a scale from 0 to 18. Among the 43 participants were 18 natives, 18 non-natives with the same L1 background than the non-native speaker and 7 non-natives with a different L1 than the non-native speaker. The results provide evidence that non-native listeners tend to trust native statements more, while native listeners showed no tendency to mistrust non-native statements. The researchers failed therefore to support the *processing fluency hypothesis*, although Podlipský et al. reported a "moderate correlation between comprehensibility and credibility of foreign-accented utterances" (2016, p. 30). It is important to notice that the design used in this experiment differed significantly to that used in Lev-Ari and Keysar (2010). To ensure that the actual truthfulness of the statements did not bias the judgment by the participants, Lev-Ari and Keysar designed a within-subject and within-item design in which speakers and statements were fully rotated across conditions. Such a rotation was not present in Podlipský et al. (2016), where a statement was always assigned to a specific speaker condition. As such, the actual credibility of statements and speaker conditions may have been confounded.

Finally, Frances et al. (2018) compared credibility ratings for different regional accents, thus testing the *processing fluency hypothesis*

in an entirely native context. A non-familiar regional accent was, similar to the non-native speech in Lev-Ari and Keysar (2010), supposed to lower processing fluency (see also Floccia et al., 2006) and hence to influence the credibility judgment of a local accent compared to a standard accent. However, their results failed to support the *processing fluency hypothesis* as statements in the regional (non-standard) accent were not to be believed to be less credible than statements that were presented with the local (standard) accent. In fact, overall stronger accented utterances were even believed to be more credible.

The potential role of accent familiarity

The seemingly contradictory findings reported above might be explained by a third factor that has not been considered so far: the familiarity of the participants with the tested accents. Some research has shown that familiarity with a specific accent facilitates its comprehension (e.g., Gass & Varonis, 1984; Derwing & Munro, 1997, Matsuura, Chiba, & Fuijeda, 1999). We therefore argue that long-term familiarity with a particular accent can increase processing fluency, in turn affecting the credibility judgements of participants. In other words, even a foreign accent, when it is highly familiar, does not necessarily disturb processing fluency. Therefore, there might be a difference in the credibility ratings depending on whether the foreign accent is familiar or not. Lev-Ari and Keysar (2010) indicated that participants from their experiments could not identify the mother tongue of the non-native speakers recorded in the experiment, when asked to do so in a questionnaire administered after the experiment. This tends to indicate that the accents were unfamiliar to them. The situation is somewhat different for several studies who failed to replicate this effect. For example, Stocker (2017) used German, Italian, English and French accents, which all represent highly habitual non-native accents for their French and German speaking participants in Switzerland, as all these languages are widely spoken in the area. In fact, Stocker (2017: 626) did acknowledge this possibility.

Similarly, a familiarity bias cannot be excluded in De Meo (2011), who used a Chinese accent in Italian. Given the increasing number of Chinese migrants in Italy since the years 1990 (Marsden, 2014), this accent may well be habitual to native Italian speakers. In the study by Souza and Markman (2013), it remains unclear whether their participants were familiarized with the tested non-native accents or not, so the bias cannot be ruled out. Among the studies that failed to reproduce the difference in credibility, only Podlipský et al. (2016) explicitly excluded the bias of familiarity, by reporting that the participants had little or no experience with the tested accent. However, the other methodological differences between their study and the others mentioned above could also have influenced the outcome of their experiment.

In a nutshell, the familiarity of the participants with the tested accents could potentially represent an important bias in previous studies and explain the differences of outcome between them. As the role of this potential bias has not been tested so far, we ran an experiment with undergraduate students of the University of Fribourg (Switzerland) in order to determine if long-term

familiarity with an accent would facilitate the processing fluency of a statement presented in this accent and, consequently, would cease to bias the credibility judgements. This experiment will also allow us to confirm or rule out the impact of a factor that has so far been neglected. In addition, by including both a familiar and unfamiliar native accent, we will also be able to assess further the negative results found by Frances et al. (2018) on the role of regional accents on speaker credibility.

Experiment testing the effect of accent familiarity

Participants

The experiment included 44 participants (40 females, mean age = 22, range 18–47) who were undergraduate psychology students at the University of Fribourg in Switzerland. All of the participants evaluated themselves as native French speakers, but several indicated linguistic competences in other languages, such as English or Spanish. None of the participants indicated having competences in the language corresponding to the unfamiliar accent included in the experiment, namely Finnish. As expected, the majority of our participants self-evaluated a medium to high level of competence in standard German (average level was B2 in the European framework of reference). None of the participants indicated having lived in either Québec (the unfamiliar variety of French included in the experiment) or in Finland in the last 5 years. Participants were granted course credit for participation.

Materials

The experiment included 24 statements presenting trivia facts, such as "Les fourmis ne dorment pas" ('Ants don't sleep'). Among them, 12 were true, and 12 were false. The items were taken from the list used in Stocker (2017), who had in turn translated and adopted the materiel from the statement list used in Souza and Markmann (2013) and Lev-Ari and Keysar (2010). The same list was used in order to maximize comparability between experiments. Lexically, the statements chosen did not give indications on the source of any of the accents included in the experiment. For this reason, sentences like "Le premier zoo des Etats-Unis a ouvert à Boston en 1876 " ('The first zoo of the United States opened in Boston in 1876') were removed, as the numbers are differently expressed depending on the linguistic variety (in Fribourg region: "mille-huit-cent-septante-six" and in Canada: "mille-huit-cent-soixante-seize"). The items were audio recorded by four male speakers between the age of 20 to 30, representative of four different accents, two of which were native and two were non-natives; two were familiar and two were unfamiliar, as summarized in Table 1. Fribourg, were the participants were recruited, is a bilingual city at the border of the German and French speaking parts of Switzerland. As a result, a Swiss German-accent in French can be considered a highly familiar accent in that city. On the other hand, according to the canton of Fribourg (Service de la statistique du

Table 1
Types of accents included in the experiment

	Native	*Familiar*
French: Swiss variety	+	+
French: Canadian variety	+	-
Swiss German	-	+
Finnish	-	-

canton de Fribourg, 2020[2]), Canadian and Finnish citizens a very rare in that region, which means that those accents can be considered to be unfamiliar for the local residents.

When choosing the speakers for the recording, care was taken to ensure that all recordings of non-native speakers were intelligible and had about the same degree of accent. The duration of the recordings was between $200 - 300$ *ms* and were read at a similar pace between the different speakers and statements. In addition, in order to ensure that the judgements were not biased by the veracity of the statements, four different lists were created using a Latin square design, so that participants only heard one version of each statement. Within each list, three of the statements uttered by a speaker were false, and three of them were true. Participants were assigned to one of the four lists randomly, and within each list, statements were presented randomly.

Procedure

The experiment was conducted online using the Qualtrics® software (Qualtrics LLC, Provo, Utah, USA). Participants were instructed to do the experiment in a quiet and non-distracting environment and asked not to search the information presented in the statements online before answering. Furthermore, they were told that the experiment was designed to test their intuitions on various information. In order to avoid the bias of stereotypes, it was indicated twice that the statements did not necessarily reflect the opinion of the speakers. Thus, as in Lev-Ari and Keysar (2010), the speakers were presented as mere messengers of the content. In order to explain to the participants why statements would be presented in a variety of accents without raising their attention to the importance of this aspect of the experiment, participants were told that the experiment had been designed by an international group of students from the University of Bern.

Before starting the actual experiment, two recordings of the Swiss-French speaker were presented in order to familiarize participants with the task. Participants were asked to judge the credibility of each recording by moving a slider on a continuous scale: [I am certain that this is false] − [I think this is false] − [I think is right] − [I am certain this is right] with regular spacing. Each recording was presented in isolation, and the participants had the possibility to replay the recordings if needed. In order to move on to the next item by clicking on a "continue" button,

the slider had to be moved. After completing the experiment, participants answered a questionnaire gathering personal data as well as completed three additional control tasks to ensure the homogeneity of our sample, in both the way they did the task and their representations of the accents in the experiment. The first one included again a recording of the Finnish accent and the participants were asked if they could identify it, and if so, where they would locate the origins of the speaker. The second additional task was meant to ensure that the participants paid attention throughout the experiment. During this task, the participants were asked to identify from a list of 14 statements the statements that they had already heard during the experiment (7 were statements from the experiment and 7 did were foils). The third task was meant to detect a potential presence of stereotypes. Participants were asked whether people from Québec, Finland, French-speaking Switzerland and German-speaking Switzerland were generally considered to be more or less credible. As in the experiment itself, they could move a slider from a scale from [very credible] to [not credible at all]. The whole experiment lasted approximately 20 minutes.

Results and analysis

In order to include both participants and items as random factors in all analyses and to avoid any fixed effect fallacies by separating by-participant and by-item analyses (Clark, 1973; Brysbaert, 2007), data were analyzed by fitting linear mixed-effects models using the *R software*® (Development Core Team, 2010, version 3.1.2). To examine the contributions of our factors of interest, *Type of accent* and *Familiarity* were added one after the other to a model only accounting for random effects (i.e. from the simplest to the more complex one). Models were built
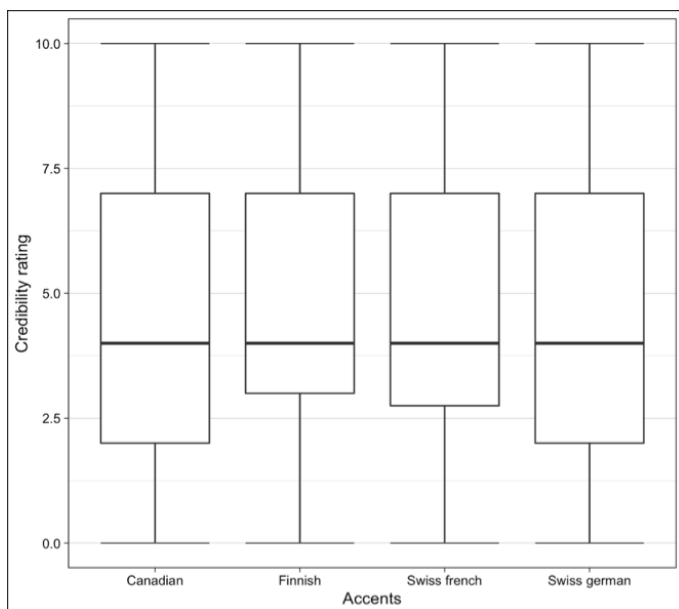

Figure 1. Credibility rating by the 44 participants for the four accent conditions on a scale to 10, (10 indicating the highest credibility)

using the *lmer()*-function of the *lmer4*-package of *R* (Kuznetsova, Bruun Brockhoff & Haubo Bojesen Christensen, 2014) and model comparisons were assessed using the *anova()*-function, which calculates the Chi-square value of the log-likelihood in order to evaluate the difference between models, following Baayen's (2008) procedure.

Our initial model – the null model – encompassed random intercepts for both *Participants* and *Items*, and a random slope for *Correctness of the item*, as participants may react very differently to the fact that some assertions were fundamentally correct, whereas others were incorrect. The inclusion of other random slopes was not possible due to convergence issues. We first compared this model to one encompassing *Type of Accent* as fixed factor, to examine claims in the literature. With comparison to our null model the model that included *Type of Accent* as a fixed factor did not show a better fit, ($\Delta\chi2 = .051$, $\Delta df = 3$, $p = .92$) (see Figure 1). We will come back to this lack of fit, as it has some important relevance for the current literature on the topic.

Additional analyses

As an effect of *Type of Accent* would have been central to our hypothesis and to support previous studies on the topic, we decided to calculate a Bayes factor on the lack of effect of *Type of Accent* to assess the relative strength of our evidence (of no effect). As such, we wanted to evaluate whether our data were sufficiently sensitive to truly support H0 (no effect) over H1 (an effect of Type of Accent) (Dienes, 2014, 2016). In order to determine the evidence for H0 over H1, a plausible range of effect size is needed. To be on the conservative side, we took the smallest (significant) effect (0.5) found in other studies using the same measure and scale (i.e., Lev-Ari & Keysar, 2010; Dragojevic & Giles, 2016; Hanzlíková & Skarnitzl, 2017) as the reference effect size. Likewise, we used a half-normal distribution to calculate our Bayes factor, to avoid favoring the probability of supporting H0 over H1 (i.e., half-Cauchy distribution; Dienes, 2019). In blunt terms, we statistically tested whether our lack of difference between our native and non-native accents truly constituted evidence for H0. To do this, we used the difference of .008 as our sample mean (Native: 4.555; Non-native: 4.547, and SE of .076 (i.e., the raw difference divided by the *t*-value given by our model encompassing Type of Accent). Using the conventional cut-off of .30 suggested by Jeffreys (1961), the resulting Bayesian analysis showed strong evidence for the null hypothesis over the existence of an effect of Type of Accent, $B = .13$. In other terms, our data was sensitive enough to truly support the null hypothesis.

Table 2
Mean and standard deviations for the conditions native, non-native, habitual and non-habitual

	Habitual		Non-habitual	
	μ	SD	μ	SD
Native	4,60	2,96	4,51	2,95
Non-native	4,51	2,93	4,59	2,96

Table 3

Model comparisons along with their statistical values, null model:
Score ~ (1|Participant)+(1|Item)+(1|Participant:Condition))

	Δχ2	Δdf	p
Score ~ Type of Accent + (1|Participant) + (1|Item) + (1|Participant:Correctness)	0.513	3	0.92
Score ~ Familiarity + (1|Participant) + (1|Item) + (1|Participant:Correctness)	0.001	1	0.97
Score ~ Native + (1|Participant) + (1|Item) + (1|Participant:Correctness)	0.010	1	0.92
Score ~ Condition + (1|Participant) + (1|Item) + (1|Participant:Correctness)	3.714	1	>.05

*Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

Results of the control tasks

In the first control task, none of the participants could identify the origins of the Finnish speaker correctly. In the second task asking them to identify the sentences that they had already heard, participants made on average 2.1 mistakes (*SD* = 1.2) meaning that they recognized on average 12.9 of 14 sentences correctly. For the third control task, we simply ran a one-way ANOVA to check whether any of the four language groups would be generally considered as more or less credible than the others. This was clearly not the case (*F*(3,172) = .137, *p* = .938, see Figure 2), demonstrating that there seem to be no stereotype associated to the general credibility of the language groups.

Discussion

Whether the lower processing fluency caused by a foreign or a regional accent affects the credibility of speakers is still very much an open issue in the literature, with only a few studies finding evidence of this effect. In this paper, we suggested a tentative explanation for the observed differences between previous studies by assessing the effect a factor that has received very little attention, namely the familiarity of the non-native accent. A foreign accent might not trigger low processing fluency if it is highly familiar to a speaker, in turn not influencing the credibility of any spoken statement. In order to assess the role of accent familiarity, we compared two native and two non-native accents in French, each accent being either highly familiar or unfamiliar.

Our results clearly indicated that familiarity was in fact not relevant to explain differences of credibility in previous work, as our participants did not rate unfamiliar variants lower compared to familiar ones. Our lack of familiarity effect mimics that of Frances et al. (2018), who also reported that a statement that is provided by a native speaker with a regional accent is not believed to be less credible than a statement presented with a standard pronunciation. In addition, we also failed to replicate the effect of non-native accents for speaker credibility put forward by Lev-Ari and Keysar (2010). Instead, we strongly supported findings from previous studies who failed to find this effect (e.g., Souza &

Markmann, 2013; Stocker, 2017). As such, our lack of effects does question the idea that the credibility of non-native speakers might be lower compared to that of native speakers. Importantly, we statistically showed that our lack of effects was not simply due to data insensitivity.

The results obtained on the control tasks also confirmed strong homogeneity of our sample. None of the participants could identify the origin of the non-native and unfamiliar accent, (i.e., the Finnish accent) correctly. Therefore, the Finish accent was truly an unfamiliar accent for our participants, contrary to the Swiss-German accent. Our second control task also confirmed that our participants were doing the task correctly, meaning that any lack of effect could not be attributed to a lack of attention. Finally, our third control task showed some homogeneity on the way our four language groups were considered, in terms of stereotype and credibility.
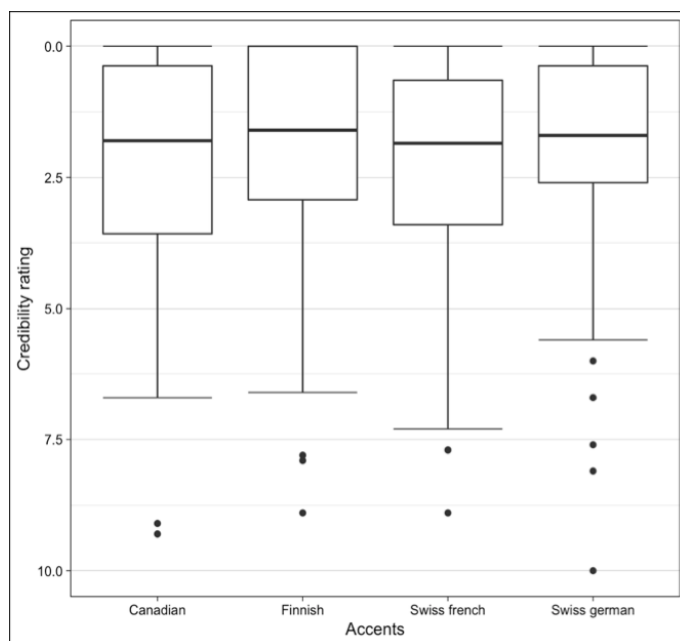

Figure 2. Estimated credibility that, according to the participants, is generally attributed to the four linguistics groups on a scale to 10 (10 indicating the lowest credibility)

An important difference with the study of Lev-Ari and Keysar (2010) is the fact that we used only one – yet different – speaker per condition. The four men came from a similar socio-economic background and within the same age range. Given the lack of effect in any of the conditions, we argue that having only one voice per condition did not have an impact on our results. We further argue that potential variations, for example in tonality and pitch (duration and volume level were constant), that could have an impact on the results may more likely depend on the language spoken than on individual difference between speakers. Järvikivi, Aalto, Aulanko, and Vainio (2007), for example, found quantity distinction in a stressed syllable in Finnish to be systematically signaled with tonal means. Although ideal, striving for an absolute equality between speakers would somehow disguise all the phonological deviations that occur when speaking a foreign language. As such, we believe that a potential bias should only be considered if the tonal variety were caused alone by the individual phonological traits of speakers, independently of their mother tongue. It is true, however, that those individual characteristics of the speakers might have had an influence on the credibility ratings provided by our participants. Even though we cannot fully exclude this bias, neither our items nor the data obtained did show any indication that this occurred. We believe that two possible alternatives may be considered in future studies on this topic to examine this issue. First, a possible way to control this potential bias would be to construct a within-speaker design, with different speakers, yet all expressing themselves in all accents. In light of the rather different sources of accents in our experiment, this would have been very difficult to implement. More generally, finding speakers that can speak homogeneously across accents may be difficult. Second, another possible way to control this potential bias may be to have different speakers per accent and per utterance, and these speakers would be accounted for in the different models' random structures. Although still quite difficult to implement, this solution may be more viable than the first one.

In all, statements that were presented by non-native speakers were not believed to be less credible that those presented by native speakers. Similarly, statements that were presented by speakers with unfamiliar accents were not believed to be less credible that those presented by speakers with familiar accents. Participants evaluated the truthfulness of spoken statements independently of the accent of speakers.

Credibility effect may still be found in other contexts than the one tested in our experiment, especially contexts in which speakers are not used to multilingualism. Indeed, the participants in our experiment were highly familiar with accented speech in general, as they live in a bilingual city, and are students used to socialize with many non-native speakers. We therefore cannot rule out that speakers who have less exposure to variations would judge statements presented with a non-native accent as less credible. For example, speakers from France, especially those speaking a variety close to the norm represented by Parisian French, are aware of speaking a variety usually considered as more correct than others (Kuiper, 1999). These speakers might also have a stronger reaction to other non-familiar accents as well as non-native ones than our

participants. Similarly, participants coming from an occupational background that does not involve regular contacts with other varieties might also be more influenced by the loss of processing fluency and might react accordingly in their judgments. In order to assess these possibilities, future work will need to include participants from other regional and/or other occupational backgrounds. Another variable that will need to be monitored more closely in future work are the actual disturbances provoked by different unfamiliar accents. Even though the Finish accent was not recognized by our participants, this accent does not necessarily have a stronger impact on the pronunciation of French phonemes compared to the German accent. It is likely that an accent that would lower processing fluency by involving stronger distortions of phonemes create stronger effects (similarly to Lev-Ari & Keysar, 2010).

Finally, another avenue of enquiry worth exploring would be the type of data that participants have to evaluate. In all the experiments reported in this paper, including ours, participants had to evaluate trivia facts were not specifically associated to them. Evaluating more self-relevant statements, or statements spoken by closer sources (e.g., evaluating a doctor's or politician's statement) might follow very different mechanisms, in turn potentially influenced by accent properties. This is reminiscent of studies suggesting that people who have a foreign accent tend to be less successful as sellers (Tsalikis, DeShields, & LaTour, 1991) and that teachers with a foreign accent receive fewer positive evaluations (Gill, 1994).

In all, the mechanisms underlying potential differences in how accents influence credibility are not yet clear. Our results provide some insight into the role of familiarity, yet mainly by indicating its rather modest influence. In addition to this, statistically, by calculating Bayes factors, we also provide strong support for the fact that statements by native speakers are perceived as credible as those by non-native ones. In applied terms, this is good news for all the people speaking a language with a non-familiar or non-native accent.

References

Alter, A. L., & Oppenheimer, D. M. (2006). Predicting short-term stock fluctuations by using processing fluency. *Proceedings of the National Academy of Sciences, 103,* 9369–9372. https://doi.org/10.1073/pnas.0601071103

Alter, A. L., & Oppenheimer, D. M. (2008). Effects of fluency on psychological distance and mental construal (or why New York is a large city, but New York is a civilized jungle). *Psychological Science, 19*(2), 161–167.

Baayen, R. H. (2008). *Analyzing linguistic data.* Cambridge University Press.

Brysbaert, M. (2007). The language-as-fixed-effect-fallacy: Some simple SPSS solutions to a complex problem. *London: Royal Holloway, University of London.*

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning*

and Verbal Behavior, 12(4), 335-359. https://doi.org/10.1016/S0022-5371(73)80014-3

De Meo, A., Vitale M., Pettorino, M., & Martin P. (2011). Acoustic-perceptual credibility correlates of news reading by native and Chinese speakers of Italian. In W.-S. Lee, E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences, City University of Hong Kong* (pp. 1366–1369). Congress of Phonetic Sciences.

De Meo, A. (2012). *How credible is a non-native speaker? Prosody and surroundings*. CLEUP. Retrieved March 19, 2020 from https://unora.unior.it/handle/11574/40089

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2014.00781

Dienes, Z. (2016). How Bayes factors change scientific practice. *Journal of Mathematical Psychology, 72,* 78–89.

Dienes, Z. (2019). How do I know what my theory predicts? *Advances in Methods and Practices in Psychological Science, 2(*4), 364–377.

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,* 396–420. https://doi.org/10.1111/hcre.12079

Dragojevic, M., Giles, H., Beck, A.-C., & Tatum, N. T. (2017). The fluency principle: Why foreign accent strength negatively biases language attitudes. *Communication Monographs, 84,* 385–405. https://doi.org/10.1080/03637751.2017.1322213

Evans, J. R., & Michael, S. W. (2014). Detecting deception in non-native English speakers. *Applied Cognitive Psychology, 28,* 226–237. https://doi.org/10.1002/acp.2990

Floccia, C., Goslin, J., Girard, F., & Konopczynski, G. (2006). Does a regional accent perturb speech processing? *Journal of Experimental Psychology: Human Perception and Performance, 32,* 1276–1293. https://doi.org/10.1037/0096-1523.32.5.1276

Gass, S., & Varonis, E. M. (1984). The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning, 34,* 65–87. https://doi.org/10.1111/j.1467-1770.1984.tb00996.x

Gill, M. M. (1994). Accent and stereotypes: Their effect on perceptions of teachers and lecture comprehension. *Journal of Applied Communication Research, 22,* 348–361. https://doi.org/10.1080/00909889409365409

Hanzlíková, D., & Skarnitzl, R. (2017). Credibility of native and non-native speakers of English revisited: Do non-native listeners feel the same? *Research in Language, 15,* 285–298. https://doi.org/10.1515/rela-2017-0016

Järvikivi, J., Aalto, D., Aulanko, R., & Vainio, M. (2007). Perception of vowel length: Tonality cues categorization even in a quantity language. In *Proceedings of the 16th International Congress of Phonetic Sciences* (ICPhS XVI) (pp. 693–696).

Jeffreys, H. (1961). *Theory of probability*. Oxford University Press.

Kuiper, L. (1999). Variation and the norm. Parisian perceptions of regional French. In D. Preston (Ed.), *Perceptual dialectology* (Vol. 1, pp. 243–262). Amsterdam: John Benjamins.

Kuznetsova, A, Bruun Brockhoff, P., & Haubo Bojesen Christensen, R. (2014). lmerTest package: Test in linear mixed effects models. *Journal of Statistical Software, 82,* 1–26.

Leach, A.-M., & Da Silva, C. S. (2013). Language proficiency and police officers' lie detection performance. *Journal of Police and Criminal Psychology, 28,* 48–53. https://doi.org/10.1007/s11896-012-9109-3

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,* 1093–1096. https://doi.org/10.1016/j.jesp.2010.05.025

Lowrey, T. M. (1998). The effects of syntactic complexity on advertising persuasiveness. *Journal of Consumer Psychology, 7,* 187–206. https://doi.org/10.1207/s15327663jcp0702_04

Marsden, A. (2014). Chinese descendants in Italy: Emergence, role and uncertain identity. *Ethnic & Racial Studies, 37,* 1239–1252. https://doi.org/10.1080/01419870.2014.883082

Matsuura, H., Chiba, R., & Fujieda, M. (1999). Intelligibility and comprehensibility of American and Irish Englishes in Japan. *World Englishes, 18,* 49–62. https://doi.org/10.1111/1467-971X.00121

Novemsky, N., Dhar, R., Schwarz, N., & Simonson, I. (2007). Preference Fluency in Choice. *Journal of Marketing Research, 44,* 347–356. https://doi.org/10.1509/jmkr.44.3.347

Oppenheimer, D. M. (2008). The secret life of fluency. *Trends in Cognitive Sciences, 12,* 237–241. https://doi.org/10.1016/j.tics.2008.02.014

Podlipský, V. J., Šimáčková, Š., & Petráž, D. (2016). Is there an interlanguage speech credibility benefit? *Topics in Linguistics, 17,* 30–44. https://doi.org/10.1515/topling-2016-0003

Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science, 9,* 45–48. https://doi.org/10.1111/1467-9280.00008

Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition, 8,* 338–342. https://doi.org/10.1006/ccog.1999.0386

Souza, A. L., & Markman, A. B. (2013). Foreign accent does not influence cognitive judgments. In *The 35th Annual Meeting of the Cognitive Science Society* (pp. 1360–1365). Austin, TX: Cognitive Science Society.