Applied Cognitive Psychology, Appl. Cognit. Psychol. **31**: 247–257 (2017)
Published online in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/acp.3322

Observers' Language Proficiencies and the Detection of Non-native Speakers' Deception

AMY-MAY LEACH[1]*, RENÉE L. SNELLINGS[1] and MARIANE GAZAILLE[2]

[1]*Faculty of Social Science and Humanities, University of Ontario Institute of Technology, Oshawa, Canada*
[2]*Département des langues modernes et de traduction, Université du Québec à Trois-Rivières, Québec, Canada*

Summary: We examined whether observers' language proficiencies affected their abilities to detect native and non-native speakers' deception. Native and non-native English speakers were videotaped as they either lied or told the truth about having cheated on a test. A total of 284 laypersons—who were either native or non-native English speakers themselves—viewed these videos and indicated whether they believed that the speakers were being truthful or deceptive. Observers were more accurate when judging native speakers than when judging non-native speakers, suggesting that perceptual fluency aided deception detection. Although there was no effect of observers' language proficiencies on discrimination, their belief that interviewees were telling the truth increased with proficiency. On the whole, these findings suggest that non-native speakers may be at greater risk of being incorrectly classified in forensic contexts. Copyright © 2017 John Wiley & Sons, Ltd.

At a major international airport, customs officers are expected to interview individuals who speak any of the world's languages (e.g., Mandarin, English, Spanish, and Arabic). Yet these officers would only be able to communicate in a subset of these languages proficiently. Audits have revealed that when interviewees and officers do not share a native language, the latter use their discretion to decide how to proceed; often, interviews are conducted in the officers' native languages (Office of the Commissioner of Official Languages, 2005). Mounting evidence suggests that this practice can affect deception (e.g., Evans, Michael, Meissner, & Brandon, 2013). We examined whether observers' knowledge of non-native speech would facilitate deception detection.

Proficiency and deception

According to Zuckerman, DePaulo, and Rosenthal's (1981) multi-factor theory, lying involves greater control, arousal, emotion and cognitive processing than telling the truth. Lie tellers may actively attempt to inhibit natural demeanor and vocal cues that could reveal the lie and display behaviors that are suggestive of honesty. Monitoring and regulating behavior, while also generating and remembering the fabrication, should create additional cognitive demands (Vrij et al., 2008a). Simultaneously, lie tellers might experience a range of emotions, such as guilt about their deception and transgression, fear of discovery, anger at being questioned, or even joy when successfully fooling the listener (e.g., Ekman, 1996). According to even the earliest theories (e.g., the James–Lange Theory of Emotion), these emotions would be associated with elevated physiological arousal.

Language proficiency also affects control, cognitive processing, arousal, and emotion. For example, when speaking in a non-native language, an individual must actively inhibit the articulatory phonetics (e.g., manner and place of articulation, voicing, and nasality of sounds) and vocabulary of the native language. Controlling these verbal behaviors should increase cognitive demands. Indeed, non-native speakers show signs of being cognitively taxed and greater neural activation (e.g., Perani & Abutalebi, 2005). Yet they can also experience emotional distance from the subject matter (Bond & Lai, 1986). Self-reports and physiological measures have revealed that insults, reprimands, and taboo words (i.e., emotional words and phrases) elicit less arousal when they are presented in a non-native (vs. native) language (Caldwell-Harris & Ayçiçeği-Dinn, 2009; Dewaele, 2004; Harris, Ayçiçeĝi, & Gleason, 2003). Thus, unlike during deception, speaking in a non-native language can inhibit emotion and physiological arousal.

Given that similar mechanisms underlie deception and non-native speech, combining these two tasks should impact interviewees significantly. Researchers are in agreement about predicted effects on cognition (e.g., Da Silva & Leach, 2013; Evans et al., 2013). If lying and speaking in a non-native language are independently cognitively taxing and require additional behavioral control, then engaging in both tasks simultaneously should significantly increase cognitive load. In fact, when individuals were asked to provide false (vs. true) alibis in their non-native language, they exhibited more signs of heightened cognitive load (e.g., offering fewer details and thinking harder) than when they were speaking in their native language (Evans & Michael, 2014; Evans et al., 2013).

It is less clear how emotions and physiological arousal would be affected by lying in a non-native language. Whereas deception is associated with heightened emotions, speaking in a non-native language reduces affect. Researchers have proposed the blunted emotion hypothesis, positing that non-native speakers' emotions would be inhibited during deception. In a direct test, however, language did not affect listeners' autonomic reactions to true and false statements (Caldwell-Harris & Ayçiçeği-Dinn, 2009). Moreover, other researchers have reported that—contrary to the blunted emotion hypothesis—only non-native English speakers appeared to be more nervous and

*Correspondence to: Amy-May Leach, Faculty of Social Science and Humanities, University of Ontario Institute of Technology, 2000 Simcoe Street North, Oshawa, Canada.
E-mail: amy.leach@uoit.ca

negative when lying than when telling the truth (Evans et al., 2013). Given these mixed findings, it is possible that the combined effects of language proficiency and veracity on emotion are subtler, and affected by motivation and context to a greater extent, than cognition.

Proficiency and deception detection

Heightened cognitive demands associated with lying and speaking in a non-native language not only affect the experience of deception but also the detection of deception. Engaging in two or more cognitively taxing tasks simultaneously can impair performance on one or both tasks (Broadbent, 1957). For example, deception is easier to detect when interviewees provide the fabricated account in reverse chronological order (Vrij, Fisher, Mann & Leal, 2008b). Because speaking in a non-native language and lying are both—independently—cognitively taxing, performing both tasks at once might similarly facilitate lie detection.

Conversely, the blunted emotion experienced by non-native speakers could impair deception detection. Observers are known to rely on cues related to emotion and arousal, such as nervousness, when detecting deceit (The Global Deception Research Team, 2006). In non-native speakers, arousal cues may be misleading: these interviewees should not exhibit increased emotionality during deception (Caldwell-Harris & Ayçiçeği-Dinn, 2009). In turn, observers would be at a disadvantage when detecting non-native speakers' deceit.

Cheng and Broadhurst (2005) were the first researchers to examine this issue empirically. They found that observers were no better able to detect the deception of interviewees who were speaking in their non-native language (i.e., English) than that of interviewees who were speaking in their native language (i.e., Cantonese). These results could suggest that the cognitive and emotional effects of speaking in a non-native language counteract each other, and as a result, language proficiency has no effect on deception detection. There are reasons to suspect that individuals did not experience the full impact of the language proficiency manipulation, however. Interviewees self-reported their language proficiencies using a single scale. It is possible that interviewees' objective language skills were similar in the native and non-native languages (i.e., the groups did not differ in terms of proficiency from the outset). In addition, several interviewees code-switched (i.e., they alternated between Cantonese and English when providing responses). This practice might have reduced the cognitive demands that would typically be associated with speaking exclusively in a non-native language, resulting in null effects.

Observers were equally able to discriminate between lie tellers and truth tellers who spoke in their native and non-native languages (i.e., Spanish vs. English) in a more recent study (Castillo, Tyson, & Mallard, 2014). The replication of Cheng and Broadhurst's (2005) findings could suggest that lie detection performance is unaffected by language proficiency. Yet similar alternative explanations exist for both sets of results. Because the researchers did not objectively assess language proficiency, they could not establish the extent to which non-native and native language skills differed. It is also possible that code-switching—which was not explicitly prohibited in the Castillo et al. (2014) study—occurred, alleviating cognitive demands in non-native speakers.

Differences between proficiency groups have emerged in other studies in which code-switching was not permitted and language proficiency was assessed. In one study, individuals provided true or false alibis entirely in English (i.e., their native or non-native language; Evans et al., 2013). Using the Psychologically Based Credibility Assessment Tool, observers detected behavioral differences indicative of heightened cognitive load (e.g., quantity of auditory and spatial details) between the proficiency groups. Notably, observers were more accurate when judging non-native (vs. native) speakers' deception.

Curiously, other researchers have reported the opposite pattern of results. Across two studies, using the same stimuli, laypersons and police officers were less accurate at detecting deception when they judged non-native English speakers than native English speakers (Da Silva & Leach, 2013; Leach & Da Silva, 2013). These observers exhibited a truth bias toward native speakers (i.e., native speakers were more likely to be judged as telling the truth)—a finding that is often reported in the lie detection literature (Bond & DePaulo, 2006)—but not non-native speakers.

Theoretical perspectives

There are several explanations for these findings. Specifically, observers' decision-making might have been affected by (a) a lack of familiarity with non-native speakers, (b) difficulties processing non-native speech, and/or (c) ingroup favoritism. We review the literature on familiarity, processing fluency, and group membership effects and explore its application to the detection of non-native speakers' deception, below.

Familiarity

Observers' exposure to non-native speakers could underlie previous studies' mixed findings. When individuals speak in a non-native language, they display more composure cues (e.g., nervousness and head or body movements; Gregersen, 2005) and provide less detailed accounts (Cheng & Broadhurst, 2005). Lie tellers are also thought to exhibit some of these behaviors (see Vrij, 2008 for a full review). Observers in Da Silva and Leach's (2013) and Leach and Da Silva's (2013) studies were primarily native English speakers within a predominately English community. They might have misattributed normal aspects of non-native speech to lying, negating the deception detection facilitation effects associated with cognitive overload. Participants in the Evans et al. (2013) study might have been more familiar with non-native speakers generally, or the exact proficiency level of the non-native speakers being interviewed specifically, than those in the Da Silva and Leach studies and, thus, avoided making this error. However, the validity of this explanation could not be tested using the existing data: neither set of researchers objectively measured linguistic familiarity, preventing comparisons between the two samples.

In general, familiarity effects on deception detection have been inconsistent (Comadena, 1982; Feeley, deTurck, & Young, 1995; McCormick & Parks, 1986). High (vs. low) geographical familiarity substantially improved deception detection by increasing observers' reliance on diagnostic verbal content (Reinhard, Sporer, & Sharmach, 2013). Yet it was a U-shaped function that best described the relationship between interpersonal familiarity and deception detection (Feeley et al., 1995). Regardless of the exact pattern of results, previous research does suggest that certain types of familiarity can facilitate deception detection.

It remains unclear whether these findings generalize to linguistic familiarity. Evans and Michael (2014) reasoned that observers from communities with a high concentration of Hispanic residents would be familiar with Hispanic non-native English speakers and better able to detect their deception, as a result. Although these observers were more likely to indicate that non-native speakers were lying than native English speakers, community exposure did not affect the relationship between proficiency and deception detection. These findings appear to suggest that linguistic familiarity cannot account for observers' superior detection of non-native speakers' deception in previous research [i.e., the pattern of results reported by Evans et al. (2013)].

The linguistic familiarity hypothesis cannot be discounted on the basis of this single study, however. Evans and Michael's (2014) operational definition of familiarity was based on two factors: (1) the ethnicities present in the communities from which they sampled (eastern vs. western Texas) and (2) the proportion of the sample that identified as Hispanic (0% vs. 80%). Yet ethnicity is not perfectly correlated with language or familiarity with a particular level of language proficiency. Hispanic communities in the USA might be more likely to contain non-native English speakers, but they are still populated by a significant number of native English speakers. Even if western Texas has a high concentration of non-native speakers, it is unclear that Evans and Michael's (2014) sample of university students was born and raised in that region and regularly interacted with community members. Because the researchers did not assess contact or the observers' own language proficiencies, even those who self-identified as Hispanic might have been primarily native English speakers who were largely unfamiliar with non-native speakers' typical behaviors. Thus, the researchers' inherent assumption that demographic characteristics were indicative of familiarity, albeit reasonable, prevents the outright rejection of the linguistic familiarity hypothesis without further testing.

Group membership

There is reason to believe that an observer's status as a non-native speaker could also affect deception detection. It is well established in the social psychological literature that observers favor their ingroup and discriminate against members of the outgroup (e.g. Tajfel & Turner, 1979). In one of the few studies that has applied this phenomenon to language groups, Guimond and Palmer's (1993) survey of college and university students revealed that unilingual English speakers exhibited ingroup favoritism (i.e., they made more negative evaluations of French speakers than English speakers), whereas bilinguals did not. Decreases in respondents' proficiency in French—the non-native language—were associated with greater ingroup bias. Thus, previous findings on the detection of non-native speakers' deception could be reframed: native English observers displayed a positive truth bias toward the ingroup (i.e., other native English speakers) and a negative lie bias toward the outgroup (i.e., non-native English speakers).

Observers who are non-native speakers themselves might exhibit one of several patterns of bias. Increases in observers' bilingualism could eliminate any differences in bias (a sign of ingroup favoritism; e.g., Guimond & Palmer, 1993). Yet there have been reported cases in which individuals who were in the linguistic minority (i.e., French native speakers living in a country in which English was the predominant language) considered native English speakers members of the outgroup (Genesee & Holobow, 1989). It is unclear whether other non-native speakers (i.e., those who were native speakers of a language other than French) would be categorized as outgroup members (because of their different language backgrounds and accents) or ingroup members (because of their membership in the broader linguistic community of non-native speakers). The latter is likely as recent work suggests that individuals overcome any distance and alienation that they feel as minority language speakers by identifying with other individuals who have nonstandard accents (Giles & Rakic, 2014). If that is the case, Tajfel and Turner's (1979) theory would predict that native English speakers (the outgroup) would be viewed more negatively than other non-native English speakers (the ingroup) by native French speakers. The limited research on the topic has failed to support this hypothesis; however, when tested, native French speakers were actually neutral toward English speakers (i.e., they were equally positive toward ingroup and outgroup speakers; Genesee & Holobow, 1989). These findings could suggest that these non-native speakers would similarly fail to exhibit any biases in terms of lie detection.

Processing fluency

Another explanation for the previous set of deception detection findings is that native speakers might have had difficulty processing non-native speech. In one study, observers were less likely to believe trivia communicated by individuals who had mild or heavy accents (Lev-Ari & Keysar, 2010). Because they were informed that these individuals were merely relaying a native speaker's statements, it is unlikely that discrimination was entirely responsible for this finding. Rather, difficulties understanding the non-native speakers might have undermined credibility judgments. Across other domains, fluency—the feeling of ease or difficulty associated with processing information—affects perceptions of risk, liking, and truthfulness (McGlone & Tofighbakhsh, 2000; Reber & Schwarz, 1999; Song & Schwarz, 2009). Indeed, auditory fluency guides even 4- and 5-year-olds' decisions about whom to believe (Bernard, Proust, & Clément, 2014). When the task is difficult, observers are not only better able to discriminate between true and false statements that are fluent (vs. dysfluent), but they also exhibit a truth bias toward fluent statements and a lie bias toward dysfluent statements (Unkelbach, 2007). These findings echo previous research examining the detection of beginner (dysfluent) English speakers' and native (fluent) English speakers' deception (Da Silva & Leach, 2013).

Research in other fields suggests that fluency is a subjective experience: the duration and frequency of previous exposure can affect decision-making (Oppenheimer, 2008). Personal experience lacking proficiency in a language could make it feel easier to process other non-native speakers' statements. Proficiency-based nonverbal behaviors (e.g., body movements; Gregersen, 2005) might be familiar, for example, as might grammatical and pronunciation errors, depending on the language of origin. Even if the messages themselves were objectively more difficult to process than those of native speakers, perceived fluency could eliminate any negative biases toward that group.

There is empirical support for an interlanguage speech intelligibility benefit in certain contexts. Perhaps unsurprisingly, native speech is more intelligible to native speakers than accented non-native speech (e.g., Bent & Bradlow, 2003; van Wijngaarden, 2001). How non-native speakers of a language perceive non-native speech is less clear. Some researchers have failed to find differences between native and non-native speakers' assessments of the intelligibility of non-native speech (e.g., van Wijngaarden, 2001). Others have reported that non-native English speakers were better able to understand high proficiency non-native speech than native speech—even when the listeners and speakers were from different linguistic backgrounds (Bent & Bradlow, 2003). Non-native speakers' subjective perceptions of the ease of understanding have also varied considerably, with members of some language groups reporting that they had less difficulty understanding non-native speakers than native speakers and others indicating the opposite (e.g., Munro, Derwing, & Morton, 2006; Smith & Bisazza, 1982). All of these studies focused on the intelligibility or comprehension of single phonemes and sentences rather than extended interviews. Whether an interlanguage speech intelligibility benefit generalizes to deception detection has not been studied directly, to date.

The present research

We explored the, at times contradictory, predictions offered by the familiarity, processing fluency, and group membership literatures about the detection of non-native speakers' deception. In this study, we examined how both observers' and speakers' language proficiencies affected deception detection. Participants with varying degrees of English proficiency assessed the veracity of native and non-native English speakers. Using signal detection analyses, we assessed participants' biases and abilities to discriminate between lie tellers and truth tellers. Thus, this work applied established social psychological theories to a novel context and added to the nascent, mixed literature on non-native speakers' deception.

Hypotheses

Discrimination

We predicted that the most proficient observers (i.e., native English speakers) would be better able to discriminate between lies and truths when judging native English speakers than when judging non-native speakers, replicating previous work (Da Silva & Leach, 2013; Leach & Da Silva, 2013). Less proficient observers were expected to be familiar with both native and non-native English speech, because of their own experiences speaking a non-native language and interacting with the broader English-speaking community, rendering discrimination similar across those two conditions. We conducted exploratory analyses to determine whether observers' global proficiency or proficiency in a specific domain (e.g., speaking proficiency) could account for these effects.

Bias

As in previous studies (e.g., Da Silva & Leach, 2013; Evans & Michael, 2014), we hypothesized that native English speakers would view non-native speakers less positively than native speakers. Other literatures suggested competing predictions for the tendencies of less proficient observers. It was possible that increases in observers' bilingualism could reduce bias toward both native and non-native speakers, in keeping with Guimond and Palmer's (1993) findings. Less proficient observers could have exhibited a truth bias toward native speakers (because of processing fluency) or a lie bias (because of outgroup bias). The ingroup–outgroup and psycholinguistics literatures suggested that observers who did not speak English as a native language could exhibit a truth bias or neutrality toward other non-native speakers. We explored all possibilities.

METHOD

Participants

In total, 284 undergraduate students participated in this study. Native English speakers (N = 68) were from an English language university; they were awarded extra credit in exchange for their participation. We excluded nine of these individuals from analyses because, although they self-identified as native English speakers, they performed at the 'basic' or 'immediate' level on at least one of the language proficiency tasks. Thus, the total sample size of native English speakers was 59 (43 women and 16 men; M age = 21.32 years, SD = 4.86). Using established [blinded] census categories, 44 participants self-identified as White, 1 as Aboriginal, 5 as Black, 1 as Filipino, 2 as Spanish, 2 as South Asian, and 4 as 'Other'. Two hundred and sixteen non-native speakers (139 women and 77 men; M age = 23.94 years, SD = 5.59), whose English-language proficiency ranged from basic to advanced, were recruited from French language and English language universities and received extra credit or 20 dollars for their participation. Using census categories, 143 of these participants self-identified as White, 3 as Aboriginal, 5 as Arab, 51 as Black, 2 as Chinese, 1 as Spanish, 1 as Japanese, 1 as Latino-

American, 1 as South Asian, 1 as South East Asian, and 4 as 'Other'. Three participants did not indicate their races.

Research design

We used a mixed-factors research design. The independent variables were speaker proficiency (i.e., native speakers vs. non-native speakers), observer proficiency (i.e., the continuous range of global proficiency scores), and veracity (lie tellers vs. truth tellers), with the last factor being within participants. The native language of observers was either English or French, whereas speakers' native languages varied (e.g., English, Arabic, and Chinese).

Materials

Video footage

We used the same video footage as two previous studies on the topic (see Da Silva & Leach, 2013 and also Leach & Da Silva, 2013 for full descriptions). We conducted a modified version of the Russano cheating paradigm (Russano, Meissner, Narchet, & Kassin, 2005). Interviewees were enticed to cheat on a test (or not) by a confederate. An experimenter, who was blind to condition, then confronted them about having similar, incorrect answers. Interviewees were informed that the integrity of the study had been compromised, and it might be considered a case of misconduct. Then, they were asked a series of closed-ended and open-ended questions in English about whether he or she had cheated on the test (e.g., 'Can you describe everything that happened from the minute I left until I returned?'; 'Did you cheat on the test?'). All interviewees included in the study—truthfully or deceptively—denied having cheated on the test.

The entire interview was captured using a hidden video camera. In total, clips of 30 interviewees (*M age* = 20.31 years, *SD* = 1.85; *M* length per clip = 92.73 s, *SD* = 32.17) were compiled. Half of the video clips (10 truth tellers and 5 lie tellers) were of native English speakers who had been recruited from an Introductory Psychology course at [blinded] University. Several different racial groups were represented: Black (n = 1), Chinese (n = 1), South Asian (n = 2), South East Asian (n = 2), White (n = 8), and Latin American (n = 1). The remaining 15 clips (10 truth-tellers and 5 lie-tellers) featured non-native English speakers who had been recruited from an English as a Second Language (ESL) school on the university campus. Their English language proficiency was assessed by the ESL school using standardized measures (Centre for Canadian Language Benchmarks, 2010); they had been classified as having 'basic' abilities (i.e., the lowest level of proficiency). Interviewees self-identified as Arab (n = 12; native language = Arabic) or Chinese (n = 3; native language = Chinese, Mandarin, or Cantonese). Where possible, truth tellers and lie tellers were yoked in terms of age, race, and gender.

In each speaker proficiency condition, clips were presented in one of two randomized orders. Specifically, there were four video sets: Native English Order A, Native English Order B, Non-native English Order A, and Non-native English Order B.

Deception detection measure

Observers were asked whether they believed that the individual in the video was lying or telling the truth (i.e., they completed a forced-choice veracity measure). They also rated their confidence in each decision on a scale from 0 (*not at all confident*) to 100 (*extremely confident*).

Deception Detection Experience Questionnaire

Deception detection experience was assessed by observers' responses to the following questions: 'Have you ever taken a course in lie detection?'; 'Have you ever participated in a study that involved lie detection?'; 'Have you ever worked in law enforcement?' Participants rated their confidence in their ability to detection deception, as well as their deception detection experience, using a 10-point scale, from 1 (*not at all confident/experienced*) to 10 (*very confident/experienced*).

Language History Questionnaire

The 13-question Language History Questionnaire asked observers to list all of the languages they spoke, how often they communicated in those languages—from 1 (daily) to 8 (less than once or twice a year)—and the ages that they learned each language. Observers rated their reading, writing, speaking, and listening proficiencies for each language on a 10-point scale, from 1 [*not at all* (*proficient*)] to 10 [*very* (*proficient*)], and how often they interacted with individuals who were speaking in their non-native language, from 1 (never) to 5 (always). Because establishing English proficiency and status as a native or non-native English speaker was paramount, observers were asked to rate both their English and French proficiencies on 5-point scales, from 1 (*poor*) to 5 (*excellent*). In addition, they were asked how many years they had spoken English, which language they considered their native language, which language they learned first, and which language they spoke at home. For each observer, we subtracted the latter from his or her current age to determine whether English had been learned within the sensitive period for language (i.e., before the age of seven; see Long, 1990, for a discussion).

English proficiency assessment

Observers were asked to complete reading, listening, writing, and speaking proficiency measures that were developed using the Centre for Canadian Language Benchmarks (2010). Reading, listening, and writing measures were administered in small groups, whereas the speaking proficiency measure was assessed individually (i.e., an experimenter tested observers one-at-a-time in a separate room). Each observer was awarded a score for reading, writing, listening, and speaking abilities (Table 1). Scores on each measure were transformed into z scores, so that they were on the same scale, and computed separately. Then, we average the z scores on the four measures to provide an assessment of global proficiency, which was used in the primary analyses.

The reading proficiency task involved answering three multiple-choice questions (e.g., 'What is the purpose of this text?') after reading a vignette. The listening proficiency task

Table 1. Distribution of observers' English language abilities by proficiency measure

Proficiency measure	Native	Advanced	Intermediate	Basic
Reading task[a]	60	144	54	18
Speaking task	60	50	74	92
Listening task[a]	60	54	72	90
Writing task	60	41	83	91

[a]*Note*: Although averaged *z*-scores were used in the actual regression analyses, for descriptive purposes and to facilitate visual comparisons between tasks, we have transformed the raw scores into a 4-point scale (i.e., 0 or 1 = *basic*, 2 = *intermediate*, 3 = *advanced*, 4 = *native*).

was similar, except that observers were asked to listen to an audio recording of the vignette. They were awarded a score of 0, 1, 2, or 3, on the basis of the number of correct answers selected. Unilingual English speakers were awarded a score of 4 to differentiate them from advanced non-native speakers.

To assess writing and speaking proficiency, observers were asked to provide a written description of one image and a verbal description of a different image. The black and white images were taken from the Thematic Apperception Test (Murray, 1943). They were open to interpretation, allowing for lengthy descriptions. Using the criteria established by the Centre for Canadian Language Benchmarks (2010), descriptions were given a score of one (basic), two (intermediate), or three (advanced). If they were provided in French, observers did not receive a score.[1] Observers were given a score of 4 when they were identified as unilingual English speakers. All descriptions were evaluated by a primary coder and secondary coder to ensure inter-rater reliability; discrepancies were resolved through discussion. Analyses indicated slightly less than a moderate level of agreement for writing proficiency, $k = .432$, $p < .001$, and a good level of agreement for speaking proficiency, $k = .707$, $p < .001$.

Procedure

Participants were greeted by an experimenter and directed to take a seat in a medium-sized classroom. They were randomly assigned to watch videos of either native or non-native English speakers. The videos were projected on a screen using MEDIALAB (Jarvis, 2006), a research software program. After viewing each video, observers completed the Deception Detection Measure. Once observers had viewed all of the videos, they were asked to complete the English Proficiency Assessment. Then, they were asked to fill out the Deception Detection Experience Questionnaire and Language History Questionnaire. Each session lasted approximately 1 hour.

RESULTS

Preliminary analyses included clip order, gender, race, experience with deception detection, number of languages spoken, and proficiency in languages other than English or French. Their inclusion did not yield significant effects; thus, reported analyses are collapsed across these factors.

Observer language proficiency was calculated using objective, standardized tests, and self-reports (i.e., observers completed a Language History Questionnaire, and we administered a Language Proficiency Assessment). We conducted a series of exploratory analyses using these different measures of English proficiency. Exploratory analyses revealed that the pattern of findings was similar regardless of how language proficiency was assessed. Specifically, bias results were consistent across all tasks, whereas findings related to discrimination varied on only one measure (i.e., the writing task). For the sake of brevity, only analyses using the global proficiency scores from the Language Proficiency Assessment are reported in full here. Analyses for each language measure within the assessment are provided in Tables 2 and 3.

Accuracy

Although unrelated to the primary hypotheses, we will briefly discuss the accuracy data here for interested readers. Correct judgments (e.g., indicating that a lie teller was lying) received a score of 1 and incorrect judgments (e.g., mistaking a lie teller for a truth teller) received a score of 0. Then, we calculated the mean truth and lie accuracy score for each observer. Overall accuracy scores ranged from a minimum of 0.13 to a maximum of 0.93 ($M = 0.54$, $SD = 0.16$).

Hierarchical multiple regression was used to assess the extent to which Speaker Proficiency and Observer Proficiency predicted accuracy at detecting lie tellers. Speaker Proficiency and Observer Proficiency were entered

Table 2. Results of a Target × Observer proficiency ANOVA on discrimination and bias by proficiency measure

	Discrimination (d')			Bias (β)		
Proficiency measure	F	p	η_p^2	F	p	η_p^2
Reading task[a]						
Target proficiency	41.18	.000	.134	6.16	.014	.023
Observer proficiency	0.16	.926	.002	0.62	.601	.007
Target × Observer interaction	1.68	.171	.019	1.02	.383	.011
Speaking task						
Target proficiency	84.76	.000	.241	10.95	.001	.039
Observer proficiency	0.21	.890	.002	1.68	.173	.018
Target × Observer interaction	1.34	.261	.015	0.33	.801	.004
Listening task[a]						
Target proficiency	88.12	.000	.248	11.36	.001	.041
Observer proficiency	0.30	.829	.003	0.67	.572	.007
Target × Observer interaction	1.91	.128	.021	1.27	.285	.014
Writing task						
Target proficiency	75.24	.000	.274	9.04	.003	.033
Observer proficiency	0.09	.964	.001	2.25	.082	.025
Target × Observer interaction	2.71	.046	.030	1.32	.269	.015

[a]*Note:* Raw proficiency scores were transformed into a 4-point scale (i.e., 0 or 1 = *basic*, 2 = *intermediate*, 3 = *advanced*, 4 = *native*).

[1] Preliminary analyses revealed that the results did not change when a score of '0' was awarded.

Table 3. Discrimination and bias of observers in each proficiency category by proficiency measure

Proficiency measure	Native speakers		Non-native speakers		Overall targets	
	d' (*SD*)	β (*SD*)	d' (*SD*)	β (*SD*)	d' (*SD*)	β (*SD*)
Reading task[a]						
Native	0.34 (0.57)	0.97 (0.43)	−0.08 (0.63)	1.16 (0.44)	0.12 (0.63)	1.07 (0.44)
Advanced	0.51 (0.47)	1.04 (0.34)	−0.27 (0.58)	1.22 (0.42)	0.10 (0.66)	1.13 (0.39)
Intermediate	0.39 (0.57)	1.16 (0.43)	−0.20 (0.58)	1.14 (0.34)	0.11 (0.65)	1.15 (0.38)
Basic	0.26 (0.64)	1.05 (0.43)	−0.20 (0.64)	1.33 (0.56)	0.03 (0.67)	1.19 (0.41)
Overall	0.43_a (0.53)	1.05_a (0.38)	-0.21_b (0.59)	1.20_b (0.42)	0.10 (0.65)	1.13 (0.41)
Speaking task						
Native	0.34 (0.57)	0.97 (0.43)	−0.08 (0.63)	1.16 (0.44)	0.12 (0.63)	1.07 (0.44)
Advanced	0.48 (0.49)	0.97 (0.32)	−0.37 (0.56)	1.18 (0.46)	0.01 (0.68)	1.09 (0.42)
Intermediate	0.43 (0.46)	1.02 (0.32)	−0.21 (0.56)	1.19 (0.35)	0.10 (0.60)	1.11 (0.35)
Basic	0.47 (0.56)	1.15 (0.41)	−0.20 (0.61)	1.24 (0.43)	0.16 (0.67)	1.19 (0.42)
Overall	0.43_a (0.53)	1.05_a (0.38)	-0.21_b (0.59)	1.20_b (0.42)	0.10 (0.65)	1.13 (0.41)
Listening task[a]						
Native	0.34 (0.57)	0.97 (0.43)	−0.08 (0.54)	1.16 (0.44)	0.12 (0.63)	1.07 (0.44)
Advanced	0.54 (0.47)	0.96 (0.36)	−0.22 (0.55)	1.26 (0.40)	0.16 (0.72)	1.10 (0.40)
Intermediate	0.54 (0.52)	1.06 (0.39)	−0.31 (0.55)	1.22 (0.45)	0.07 (0.67)	1.15 (0.43)
Basic	0.36 (0.52)	1.14 (0.35)	−0.21 (0.54)	1.17 (0.39)	0.08 (0.59)	1.16 (0.37)
Overall	0.43_a (0.53)	1.05_a (0.38)	-0.21_b (0.59)	1.20_b (0.42)	0.10 (0.65)	1.13 (0.41)
Writing task						
Native	0.34_a (0.50)	0.97 (0.41)	-0.08_b (0.60)	1.22 (0.41)	0.12 (0.63)	1.07 (0.41)
Advanced	0.40_a (0.46)	0.97 (0.34)	-0.16_b (0.54)	1.25 (0.42)	0.14 (0.61)	1.03 (0.41)
Intermediate	0.55_a (0.49)	0.98 (0.25)	-0.37_b (0.60)	1.10 (0.40)	0.02 (0.69)	1.13 (0.33)
Basic	0.42_a (0.56)	1.19 (0.43)	-0.15_b (0.63)	1.16 (0.44)	0.15 (0.64)	1.20 (0.44)
Overall	0.43_a (0.53)	1.05_a (0.38)	-0.21_b (0.59)	1.20_b (0.42)	0.10 (0.65)	1.13 (0.41)

Note: Means sharing a common subscript were not statistically different at α = 0.05 according to post hoc analyses of significant main effects and interactions.
[a]Scores were transformed into a 4-point scale (i.e., 0 or 1 = *basic*, 2 = *intermediate*, 3 = *advanced*, 4 = *native*).

at Step 1, which explained 17.9% of the total variance in the model. After entering the interaction term (Speaker Proficiency × Observer Proficiency) into the model at Step 2, the total variance explained by the model increased slightly to 18.5%, $F(3, 274) = 20.53$, $p < .001$. The addition of the interaction explained an additional 0.6% in lie accuracy, R^2 change = 0.01, F change (1, 271) = 1.95, $p = .164$; Cohen's $f^2 = 0.00$. In the final model, Speaker Proficiency was significant ($\beta = -0.385$, $p < .001$). Observers were more accurate when judging native-language speakers than when judging non-native English speakers (Figure 1). The effect of Observer Proficiency was also significant ($\beta = -0.167$, $p = .002$). Observers were less accurate at detecting lies as their own proficiency in English increased.

We also examined whether Speaker Proficiency and Observer Proficiency predicted observer's accurate identification of truth tellers. Speaker Proficiency and Observer Proficiency, which were entered at Step 1, explained 16.0% of the total variance. This total variance explained by the model remained 16.0% once the interaction term (Speaker Proficiency × Observer Proficiency) was entered at Step 2, $F(3, 274) = 17.19$, $p < .001$. Thus, the addition of the interaction did not explain any of the variance in the model, Cohen's $f^2 = 0.00$. Speaker Proficiency ($\beta = -0.340$, $p < .001$) was significant in the final model. Observers were more accurate when judging native-language speakers than when judging non-native English speakers (Figure 1). In addition, as observers' proficiency in English increased, so did their accuracy at detecting truth tellers ($\beta = 0.217$, $p < .001$).

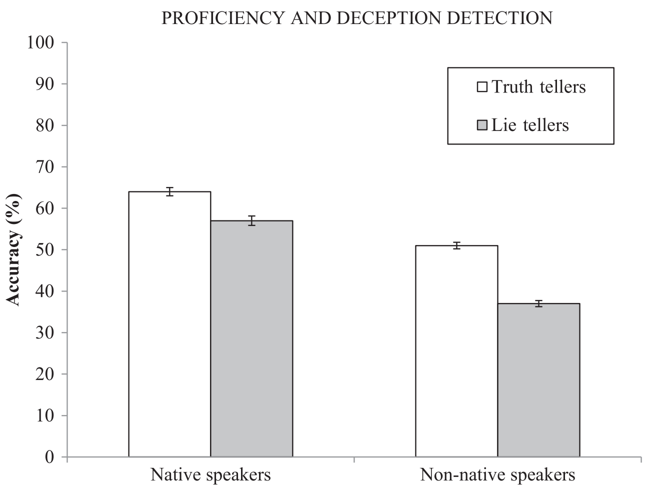

Figure 1. Mean accuracy (%) for observers in each speaker language condition. Observers were more accurate when detecting the deception of native (vs. non-native) speakers. Standard errors are represented by the error bars attached to each column

Signal detection

Meissner and Kassin (2002) have suggested that signal detection theory provides a framework for separating lie detection performance into two different parameters: discrimination and response bias. Discrimination is an observer's ability to detect a signal when it is present and reject the trial when the signal is absent. In deception detection terms, that reflects the observers' ability to differentiate between lie tellers and truth tellers. Response bias is the degree of evidence that is necessary for an observer to report that the signal has, in fact, been presented. Here, we examined whether observers displayed a tendency to indicate that individuals were lying. Correctly identifying a truth teller was deemed a 'hit', whereas incorrectly indicating that a lie teller was telling the truth was labelled a 'false alarm'. All calculations, which were based on Wixted and Lee's formula (2014), accounted for the unequal number of truth tellers and lie tellers.

Discrimination

We examined whether Speaker Proficiency and Observer Proficiency predicted discrimination (i.e., d') using hierarchical multiple regression. When these variables were entered, the total variance explained by the model was 25.0%. After entering the interaction term, 25.1% of the total variance was explained, $F(3, 274) = 30.33$, $p < .001$. Thus, the addition of the interaction did not explain any of the variance in the model, Cohen's $f^2 = 0.00$. Speaker Proficiency was the only significant factor in the final model ($\beta = -0.499$, $p < .001$). That is, overall, observers were better able to discriminate between lie tellers and truth tellers who were native English speakers ($M = 0.43$, $SD = 0.52$) than those who were non-native English speakers ($M = -0.21$, $SD = 0.59$). Using one sample *t*-tests, we compared d' values to zero (i.e., no sensitivity). Observers were able to discriminate between lie-telling and truth-telling native speakers, $t(134) = 9.54$, $p < .001$; Cohen's $d = 0.82$, and non-native speakers, $t(139) = -4.25$, $p = .000$; Cohen's $d = -0.36$.

Bias

We also conducted a hierarchical multiple regression to examine response bias (i.e., β). The two independent variables—Speaker Proficiency and Observer Proficiency—were entered first; they accounted for 5.0% of the total variance. Overall, 5.3% of the variance was explained when the interaction term was entered, $F(3, 274) = 5.08$, $p = .002$; R^2 change Cohen's $f^2 = 0.00$. In the final model, Speaker Proficiency was significant ($\beta = 0.185$, $p = .002$): non-native English speakers were labelled lie tellers more often than were native speakers. In addition, as Observer Proficiency increased, so did the belief that both native and non-native English speakers were telling the truth ($\beta = -0.128$, $p = .031$). We used one sample *t*-tests to compare observers' β values to one (i.e., no bias). Observers did not exhibit any bias when judging native speakers ($M = 1.05$, $SD = 0.38$), $t(134) = 1.51$, $p = .134$; Cohen's $d = 0.13$. However, they exhibited a lie bias toward non-native speakers ($M = 1.20$, $SD = 0.42$), $t(139) = 5.63$, $p < .001$; Cohen's $d = 0.48$.

Confidence

Finally, we used hierarchical multiple regression to assess observers' confidence in their judgments. Entering Speaker Proficiency and Observer Proficiency explained 1.5% of the variance. Adding the interaction term increased this value to 1.7%, $F(3, 274) = 1.59$, $p = .191$; R^2 change Cohen's $f^2 = 0.00$. No variables were significant in the final model.

DISCUSSION

We examined whether observers' and speakers' language proficiencies affected deception detection. Regardless of their own proficiencies, observers were more accurate when judging native than non-native English speakers. These findings were consistent with previous research (e.g., Da Silva & Leach, 2013; Leach & Da Silva, 2013).

We posited that personal experience as a non-native speaker of a language would facilitate the detection of other non-native speakers' deception. However, the data failed to support this hypothesis: the ability to discriminate between lie-telling and truth-telling non-native speakers did not vary according to observer language proficiency. This finding replicated previous exploratory work with a smaller sample and different measures of observer proficiency (Leach & Da Silva, 2013). Familiarity with non-native speech in general might not afford observers an advantage. Less proficient observers might not be in tune with their own verbal and nonverbal behaviors when speaking English, making it difficult for them to interpret these cues in others. Moreover, in previous studies, familiar observers were able to rely upon their previous experience with the critical event to aid their decisions (e.g., Reinhard et al., 2013). Having experience speaking in a non-native language would not have provided observers with comparable additional information or insight.

Similarly, processing fluency effects (i.e., the ease with which observers could process accounts) might only extend to certain lie detection contexts. Previous studies on the topic had observers rating a series of statements read by individuals who had no intent to deceive (e.g., Lev-Ari & Keysar, 2010; McGlone & Tofighbakhsh, 2000). Deception is more complex and dynamic: lie tellers are constantly monitoring and modifying, their verbal and non-verbal behaviors (e.g., Buller & Burgoon, 1996). Decision-making is also likely based on an overall assessment of accounts – that contain both true and false information – rather than a single statement. Our findings indicate that general processing fluency can account for differences in this type of higher-order discrimination only to a point: (fluent) native speakers' deception was more accurately detected than that of (dysfluent) non-native speakers. However, there did not appear to be an interlanguage speech intelligibility benefit – in terms of discrimination – for observers who were non-native speakers themselves (i.e., Speaker and Observer Proficiency did not interact).

These observers were not more likely to believe fellow non-native speakers (vs. native speakers), either. Our inability to replicate Guimond and Palmer's (1993) work on language group membership suggests that ingroup-outgroup bias was not a contributing factor here. Rather, our findings align with previous evidence that poor proficiency in a particular language does not necessarily facilitate the comprehension – and, in turn, believability – of all other non-native speakers of that language (e.g., Munro et al., 2006). There are indications that processing fluency affected observer bias in two ways, however. Observers exhibited a lie bias toward non-native speakers, but not native speakers, replicating previous research (e.g., Da Silva

& Leach, 2013). There were stark differences in the communicative abilities of the two groups: native speakers were completely fluent in English, whereas non-native speakers had just begun to learn the language. Thus, one interpretation of our findings is that the latter group's messages were more difficult to process, rendering them less believable (as observed in other domains; e.g., McGlone & Tofighbakhsh, 2000; Unkelbach, 2007). Similar logic can be applied to the finding that observers who had higher proficiencies were most likely to have a truth bias toward all interviewees. Arguably, being more proficient in English would have facilitated the processing of English accounts in general, thereby increasing message belief.

Yet this did not translate into differences in observers' confidence in their decisions. We did not have a priori hypotheses related to confidence but had explored that factor as a potential measure of processing fluency. It was conceivable that an individual who subjectively experienced the task as easy—because he or she did not have to exert additional effort trying to understand the interviewee—might have expressed higher levels of confidence. However, that did not occur. Perhaps confidence and processing fluency are related to each other but unaffected by language proficiency. Alternatively, confidence might not be an appropriate proxy for processing fluency as each factor focuses on different aspects of the decision-making process. That is, fluency is related to perceiving and understanding information (i.e., the speaker's message), whereas confidence is based on the action taken (i.e., the observer's decision to label the speaker a lie teller or truth teller). Future research should directly measure processing fluency during deception detection to test these hypotheses.

Limitations and future research

Our study emulated the experience of a border official who would have little time to decide whether a stranger was lying or telling the truth. Although exposure time, opportunities for interaction, and expertise might vary across law enforcement contexts, they should not affect the detection of deception, in general (Aamodt & Custer, 2006; Ambady & Rosenthal, 1992; Forrest & Feldman, 2000), or non-native speakers' deception, in particular. Underlying demands on the interviewee could be context-dependent. Our interview might not have been particularly cognitively taxing because it occurred immediately after the transgression, and little elaboration was required. Being accused of cheating on a test should have elicited strong emotions (e.g., guilt) in our sample, however. The blunting of emotion associated with non-native speech (e.g., Bond & Lai, 1986) could explain why observers were less able to discriminate between lie-telling and truth-telling non-native (vs. native) speakers. Telling a lie that requires additional cognitive resources (e.g., memory and attention) while performing another challenging task (e.g., providing a full narrative in a non-native language) might have the opposite effect. Indeed, researchers (Evans & Michael, 2014; Evans et al., 2013) have found different patterns of results when non-native speakers provided cognitively demanding alibis. Comparing how the proficiencies of speakers and observers interact to affect deception detection within contexts that place a greater emphasis on executive function (e.g., providing a fabricated alibi) or emotion (e.g., lying about illegal contents of one's luggage) might be an interesting avenue for researchers to pursue.

The majority of the observers in our study spoke only French and English. However, none of the interviewees was a native French speaker, and they all had heavy Arabic or Chinese accents when speaking in English. Observers who were not proficient in English may not have been familiar with—or better able to process—their speech. Thus, our findings suggest that there are limits to the interlanguage speech intelligibility benefit: Specific knowledge of a particular language or accent might be required. Although the benefit has been found under mismatched conditions (i.e., when observers and speakers were not from the same language backgrounds), it might be restricted to speakers with advanced or intermediate language proficiencies (e.g., Bent & Bradlow, 2003; Munro et al., 2006). Because non-native speakers in our study exhibited low proficiency in English, no processing advantage was afforded. Examining whether non-native observers are better able to detect the deception of high proficiency speakers across languages should be the subject of further study. So, too, should non-native English speakers and observers who share a language background. Additional research could focus on languages in which matched interlanguage speech intelligibility benefits have been established across different proficiency levels, such as Korean (Bent & Bradlow, 2003).

We inferred that interviewees who were native English speakers would form the outgroup for observers who were not proficient in English in our study because the sample was drawn from a similar region and linguistic group as in previous group identity work (e.g., Genesee & Holobow, 1989). Similarly, recent research has suggested that minority language speakers can consider other individuals with accents part of their ingroup (e.g., Giles & Rakic, 2014); we posited that the same would occur here. We did not collect data related to identity and group membership, however, and cannot completely ascertain how the interviewees were categorized by participants. For example, it is possible that observers who spoke French as a native language considered native English speakers *and* non-native English speakers (i.e., native Arabic and Chinese speakers) outgroup members. That could explain why they were less likely to believe all interviewees. Future research should establish observers' identities prior to testing, manipulate the group membership of interviewees, and explicitly assess ingroup-outgroup categorization. Our study offers a preliminary perspective on this issue, as it suggests that belonging to a broad linguistic group – non-native speakers – does not elicit typical ingroup-outgroup biases. However, further research is needed.

Although we replicated previous findings related to the detection of non-native speakers deception (e.g., Da Silva & Leach, 2013; Leach & Da Silva, 2013), these studies were not entirely independent. The same native and non-native speakers were used as stimuli in all three studies, including the current work. One could contend that the findings do not represent three separate tests of language proficiency effects and might not be generalizable to different interviewees. A solution would be to use different stimuli to replicate the current results. That does not seem necessary, however. In each study, the observers differed; using similar stimuli allowed for direct replications and extensions of previous work. More importantly, a recent study featuring non-native speakers with varying levels of language proficiency (i.e., novel stimuli) replicated the basic finding that deception is more difficult to detect in beginner English speakers than native English speakers (Elliott & Leach, 2016). That work suggests that there was nothing unique about the footage of native and non-native speakers used in the current study, and our results would generalize to different samples of interviewees who were native and non-native English speakers.

Implications and conclusions

According to the Office of the Commissioner of Official Languages (2005), individuals do not always have the opportunity to speak to justice officials in their native languages. This report is troubling because our findings suggest a lack of equivalence with respect to decision making involving native and non-native speakers. Existing practices, which allow individuals to be interviewed in their non-native languages, run counter to the United Nations' assertion that all individuals are equal under the law (UN General Assembly, 1948). Law enforcement agencies could address this issue by placing a greater emphasis on multilingualism to ensure that suspects and witnesses can be interviewed in their native languages. Of course, it is not feasible to employ officers who speak every one of the world's 7102 languages (Lewis, Simons, & Fennig, 2015). However, our research demonstrates that using interviewers who are not fluent in a language themselves do not counteract proficiency effects. It is possible that allowing individuals to speak in their native languages, through interpreters, would improve accuracy. Emerging evidence suggests that the use of interpreters does affect deception detection (e.g., Ewens et al., 2016); more research on that subject is urgently needed. In the interim, law enforcement officials should be made aware of the pitfalls associated with conducting interviews in a person's non-native language.

ACKNOWLEDGEMENTS

This research was supported by a grant to the first author from the Social Sciences and Humanities Research Council of Canada (grant no. 410-2011-1674).

REFERENCES

Aamodt, M. G., & Custer, H. (2006). Who can best catch a liar? A meta-analysis of individual differences in detecting deception. *Forensic Examiner*, *15*, 6–11.

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, *2*, 256–274. https://doi.org/10.1037/0033-2909.111.2.256.

Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *Journal of the Acoustical Society of America*, *114*, 1600–1610. https://doi.org/10.1121/1.1603234.

Bernard, S., Proust, J., & Clément, F. (2014). The medium helps the message: Early sensitivity to auditory fluency in children's endorsement of statements. *Frontiers in Psychology*, *5*, 1664–1678. https://doi.org/10.3389/fpsyg.2014.01412.

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, *10*, 214–234. https://doi.org/10.1207/s15327957pspr1003_2.

Bond, M. H., & Lai, T. M. (1986). Embarrassment and code-switching into a second language. *Journal of Social Psychology*, *126*, 179–186.

Broadbent, D. E. (1957). A mechanical model for human attention and immediate memory. *Psychological Review*, *64*, 205–215.

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal Deception Theory. *Communication Theory*, *6*, 203–242.

Caldwell-Harris, C. L., & Ayçiçeği-Dinn, A. (2009). Emotion and lying in a non-native language. *International Journal of Psychophysiology, 71,* 193–204. https://doi.org/10.1016/j.ijpsycho.2008.09.006.

Castillo, P. A., Tyson, G., & Mallard, D. (2014). An investigation of accuracy and bias in cross-cultural lie detection. *Applied Psychology in Criminal Justice*, *10*, 66–82.

Centre for Canadian Language Benchmarks. (2010). Retrieved on August 25, 2011 from http://www.language.ca/display_page.asp?page_id=1

Cheng, K. H. W., & Broadhurst, R. G. (2005). The detection of deception: The effects of first and second language on lie detection ability. *Journal of Psychiatry, Psychology and Law*, *12*, 107–118. https://doi.org/10.1375/pplt.2005.12.1.107.

Comadena, M. E. (1982). Accuracy in detecting deception: Intimate and friendship relationships. In M. Burgoon (Ed.), *Communication yearbook 6*, (pp. 446 – 472). Beverly Hills, CA: Sage.

Da Silva, C. S., & Leach, A.-M. (2013). Detecting deception in second-language speakers. *Legal and Criminological Psychology*, *18*, 115–127. https://doi.org/10.1111/j.20448333.2011.02030.x.

Dewaele, J.-M. (2004). The emotional force of swearwords and taboo words in the speech of multilinguals. *Journal of Multilingual and Multicultural Development*, *25*, 204–222.

Ekman, P. (1996). Why don't we catch liars? *Social Research*, *63*, 801–817.

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.

Evans, J. R., Michael, S. W., Meissner, C. A., & Brandon, S. E. (2013). Validating a new assessment method for deception detection: Introducing a psychologically based credibility assessment tool. *Journal of Applied Research in Memory and Cognition*, *2*, 33–41. https://doi.org/10.1016/j.jarmac.2013.02.002.

Elliott, E., & Leach, A.-M. (2016). You must be lying because I don't understand you: Language proficiency and lie detection. *Journal of Experimental Psychology: Applied*, *22*, 488–499. https://doi.org/10.1037/xap0000102.

Ewens, S., Vrij, A., Leal, S., Mann, S., Jo, E., & Fisher, R. (2016). The effect of interpreters on eliciting information, cues to deceit and rapport. *Legal and Criminological Psychology*, *21*, 286–304. https://doi.org/10.1111/lcrp.12067.

Feeley, T. H., deTurck, M. A., & Young, M. J. (1995). Baseline familiarity in lie detection. *Communication Research Reports*, *12*, 160–169. https://doi.org/10.1080/08824099509362052.

Forrest, J. A., & Feldman, R. S. (2000). Detecting deception and judge's involvement: Lower task involvement leads to better lie detection. *Personality and Social Psychology Bulletin*, *26*, 118–125. https://doi.org/10.1177/0146167200261011.

Genesee, F., & Holobow, N. E. (1989). Change and stability in intergroup perceptions. *Journal of Language and Social Psychology*, *8*, 17–38. https://doi.org/10.1177/0261927X8900800102.

Giles, H., & Rakic, T. (2014). Language attitudes: Social determinants and consequences of language variation. In T. M. Holtgraves (Ed.), *The Oxford handbook of language and social psychology*, (pp. 11 – 26). New York: Oxford University Press.

Gregersen, T. S. (2005). Nonverbal cues: Clues to the detection of foreign language anxiety. *Foreign Language Annuals*, *38*, 388–400. https://doi.org/10.1111/j.19449720.2005.tb02225.x.

Guimond, S., & Palmer, D. L. (1993). Developmental changes in ingroup favouritism among bilingual and unilingual francophone and anglophone students. *Journal of Language and Social Psychology*, *12*, 318–351.

Harris, C. L., Ayçiçeği, A., & Gleason, J. B. (2003). Taboo words and reprimands elicit greater autonomic reactivity in a first language than in a second language. *Applied PsychoLinguistics*, *24*, 561–579. https://doi.org/10.1017/S0142716403000286.

Jarvis, B. G. (2006). *MediaLab (version 2006.1) [computer software]*. New York, NY: Empirisoft Corporation.

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.

Lewis, M. P., Simons, G. F., & Fennig, C. D. (Eds) (2015). *Ethnologue: Languages of the world*, (18th edn). Dallas, TX: SIL International Retrieved from http://www.ethnologue.com.

Long, M. H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition*, *12*, 251–285. https://doi.org/10.1017/S0272263100009165.

McCormick, S. A., & Parks, M. R. (1986). Deception detection and relational development: The other side of trust. In M. L. McLaughlin (Ed.), *Communication yearbook 9*, (pp. 377 – 389). Beverly Hills, CA: Sage.

McGlone, M. S., & Tofighbakhsh, J. (2000). Birds of a feather flock conjointly(?): Rhyme as a reason in aphorisms. *Psychological Science*, *11*, 424–428. https://doi.org/10.1111/1467-9280.00282.

Meissner, C. A., & Kassin, S. M. (2002). "He's guilty!": Investigator bias in judgments of truth and deception. *Law and Human Behavior*, *26*, 469–480. https://doi.org/10.1023/A:1020278620751.

Munro, M. J., Derwing, T. M., & Morton, S. L. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition, 28,* 111–131. https://doi.org/10.10170S0272263106060049.

Murray, H. A. (1943). *Thematic apperception test*. Cambridge: Harvard University Press.

Office of the Commissioner of Official Languages (2005). Retrieved October 6, 2011, from http://www.ocol-clo.gc.ca/html/index_e.php

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.

Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current Opinion in Neurobiology*, *15*, 202–206. https://doi.org/10.1016/j.conb.2005.03.007.

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.

Reinhard, M.-A., Sporer, S. L., & Sharmach, M. (2013). Perceived familiarity with a judgmental situation improves lie detection ability. *Swiss Journal of Psychology*, *72*, 43–52. https://doi.org/10.1024/1421-0185/a000098.

Russano, M. B., Meissner, C. A., Narchet, F. M., & Kassin, S. M. (2005). Investigating true and false confession within a novel experimental paradigm. *Psychological Science*, *16*(6), 481–486. https://doi.org/10.1111/j.0956-7976.2005.01560.x.

Smith, L., & Bisazza, J. (1982). The comprehensibility of three varieties of English for college students in seven countries. *Language Learning*, *32*, 259–269.

Song, H., & Schwarz, N. (2009). If it's difficult to pronounce, it must be risky: Fluency, familiarity, and risk perception. *Pychological Science*, *20*, 135–138. https://doi.org/10.1111/j.1467-9280.2009.02267.x.

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin, & S. Worchel (Eds.), *The social psychology of intergroup relations*, (pp. 33 – 47). Monterey, CA: Brooks-Cole.

The Global Deception Research Team (2006). A world of lies. *Journal of Cross-Cultural Psychology*, *37*, 60–74. https://doi.org/10.1177/0022022105282295.

UN General Assembly, Universal Declaration of Human Rights, 10 December 1948, 217 A (III), available at: http://www.un.org/Overview/rights.html [accessed 2 December 2015]

Unkelbach, C. (2007). Reversing the truth effect: Learning the interpretation of processing fluency in judgments of truth. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 219–230. https://doi.org/10.1037/0278-7393.33.1.219.

van Wijngaarden, S. J. (2001). Intelligibility of native and non-native Dutch speech. *Speech Communication*, *35*, 103–113.

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities*, (2nd edn). New York: John Wiley & Sons.

Vrij, A., Fisher, R., Mann, S., & Leal, S. (2008b). A cognitive load approach to lie detection. *Journal of Investigative Psychology and Offender Profiling*, *5*, 39–43. https://doi.org/10.1002/jip.82.

Vrij, A., Mann, A. A., Fisher, R. P., Leal, S., Milne, R., & Bull, R. (2008a). Increasing cognitive load to facilitate lie detection: The benefit of recalling an event in reverse order. *Law and Human Behavior*, *32*, 253–265. https://doi.org/10.1007/s10979007-9103-y.

Wixted, J. T., and Lee, K. (2014). Signal Detection Theory. Retrieved from http://www.kangleelab.com/signal%20detection%20theory.html

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In L. Berkowitz (Ed.), *Advances in experimental social psychology*, (Vol. *14*, pp. 1 – 59). New York: Academic Press.

