Integrating Non-Native Speaker Identity in Semantic and Pragmatic Processing

Martin Ho Kwan Ip (mhkip@sas.upenn.edu)
Integrated Language Science and Technology Initiative (ILST)
and Departments of Linguistics and Psychology
University of Pennsylvania, 3401 Walnut St, Philadelphia, PA 19104

Anna Papafragou (anna4@sas.upenn.edu)
Integrated Language Science and Technology Initiative (ILST)
and Department of Linguistics
University of Pennsylvania, 3401 Walnut St, Philadelphia, PA 19104

Abstract

Little research to date has examined how listeners integrate cues to non-native speaker identity in real time sentence processing. Here, we examine listeners' interpretation of the semantic and socio-pragmatic content of utterances produced by either a foreign accented speaker or a native speaker. Overall, our findings suggest that processing speed was slower in the presence of foreign accents. However, the extra perceptual demands of processing unfamiliar accents did not translate into listeners' accuracy rates, and in certain sentence contexts, non-native speakers were also more likely to elicit higher semantic or pragmatic interpretation accuracy. Our findings show that non-native speaker identity plays an important role in listeners' sentence interpretations.

Keywords: socio-pragmatic inferences; sentence processing; non-native accent perception; semantic processing

Introduction

When Soviet leader Nikita Khrushchev said "*Мы вас похороним!*", or "*We will bury you!*", to Western bloc diplomats in 1956, the statement caused a frenzy in Western nations. Only decades later was it clarified that Khrushchev's exclamation was borrowed from a catchphrase during bitter exchanges between farming experts and agrarian scientists in the 1920s and 30s, which at the time were sloganized as a dispute on "who will bury whom" (Gorbachev, 1987). Rather than intending his message as a threat, what Khrushchev tried to say was that socialism would outlive capitalism: he was being figurative, not literal.

The Khrushchev case, and many others showing intergroup misunderstandings, highlight two important questions for sentence processing. First, why are utterances sometimes interpreted in a pragmatically enriched way and sometimes in a literal way? Second, does speaker identity affect utterance interpretation? To examine these questions, the present experiments investigate the effects of a speaker's foreign accent and non-native speaker identity on listeners' semantic and socio-pragmatic processing.

According to the Gricean framework, communication is a co-operative activity where listeners expect speakers to be, among other things, as informative as required given the goals of the conversation (Grice, 1975). When listeners encounter an utterance that appears to be in violation of this expectation, they have the chance to pragmatically enrich the literal meaning of the sentence. A highly cited example of this phenomenon is the case of scalar implicatures (see Noveck & Posada, 2003; Bott & Noveck, 2004; De Neys & Schaeken, 2007; Hunt et al., 2013; Tomlinson et al., 2013; Papafragou & Skordos, 2016), in which quantifiers like "some" are used instead of the stronger alternative "all" (e.g., "Mary ate some of the cakes") to convey a pragmatic meaning (i.e., upper bounded interpretation: "Mary ate some (but not all) of the cakes"), as opposed to a literal semantic meaning (i.e., lower bounded interpretation: "Mary ate some (and possibly all) of the cakes").

There are still ongoing debates as to how listeners derive scalar implicatures, and conflicting experimental evidence exists as to how scalar terms are computed (for a review, see Breheny, 2019), even in studies using the same paradigms (e.g., Grodner, Klein, Carbary, & Tanenhaus, 2010; Huang & Snedeker, 2009). One line of research inspired by these debates has found that epistemic properties of individual speakers can affect whether listeners compute utterances in a literal vs. pragmatic way. Bergen and Grodner (2012), for example, showed that individuals do not process underinformative 'some' sentences in a pragmatic way when they are led to believe that the speaker was not knowledgeable: when processing a sentence such as "Some of the real estate investments lost money", participants were less likely to infer that "some (but not all) of the money was lost" after they were told that they speaker did not meticulously compile the accounts. Similarly, listeners can also make speaker-specific adaptations to process sentences with "some" and "many" in reference to various quantities of objects (Yildirim, Degen, Tanenhaus, & Jaeger, 2015). These findings show that awareness of the speaker's knowledge state can modulate how listeners interpret the language input.

A less studied issue, however, is whether listeners also integrate cues signaling the *social identity* of the speaker, including traits that are stable across contexts (e.g., gender, native language background). In recent years, studies have shown that listeners perceive and interpret sentences differently depending on whether the speaker speak with a foreign or a native accent. For example, P600 neural responses to syntactic errors (e.g., "She mow the lawn") are attenuated when the errors are made by a non-native speaker (Hanulíková, van Alphen, van Goch, & Weber, 2012). Likewise, implausible utterances with syntactic errors (e.g., "The girl was kicked by the ball") are more likely to be interpreted in a plausible way when delivered in a foreign accent compared to a native accent (Gibson et al., 2017).

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society.*

These findings indicate that listeners have different underlying expectations for non-native compared to native speakers, and the processing effects derived from these expectations are immediate. Since foreign accented speakers are generally expected to be less linguistically competent compared to native speakers, listeners are more likely to ignore these language errors when they are produced by non-native speakers. Consistent with this view, research has also shown that listeners process foreign accented speech in a less detailed way (e.g., Lev-Ari & Keysar, 2012), and trivia statements (e.g., "Ants don't sleep") produced with a foreign accent are rated as less credible (Lev-Ari & Keysar, 2010).

In the domain of socio-pragmatic processing, different speaker expectations can also affect how listeners interpret the communicative intentions of foreign accented vs. native speakers. In a study examining under-informativeness in native and non-native speakers (Fairchild, Mathis, & Papafragou, 2020), readers were given explicit information about a character's language background (e.g., "Emma is from Boston and has a strong Boston accent" vs. "Yuqi is from China and has a strong Chinese accent") and then read a story in which the character failed to offer complete information to a friend (e.g., said that "there were apples and bananas" in the fridge but did not mention that there were also pears). When asked about the reasons for the omission, readers invoked incompetence more often for the non-native than for the native speaker (e.g., "She didn't know the word for pears"); inversely, they invoked unwillingness more often for the native speaker (e.g., "She wanted to keep the pears"). In subsequent work (Ip & Papafragou, 2021), this effect has been extended to spoken utterances and more extreme scenarios where omitting information would normally be considered deceptive: in such cases, listeners are less likely to rate underinformative foreign accented speakers as dishonest compared to underinformative native speakers.

In particular, because non-native speakers are expected to be less linguistically competent, readers have also been found to rate underinformative sentences more highly when they are attributed to a non-native compared to a native speaker – a 'pragmatic lenience' effect. In Fairchild and Papafragou (2018), readers rated sentences involving everyday common knowledge. Fairchild and Papafragou created four different types of sentences involving the quantifiers "all" or "some"; (1) true "all" sentences (e.g., "All snow is cold and can melt"), (2) false "all" sentences (e.g., "All women are doctors"), (3) semantically true "some" sentences carrying a false "not all" implicature (e.g., "Some elephants have trunks", henceforth false-"some"), and (4) true and felicitous "some" sentences (e.g., "Some people have dogs as pets"). Overall, true sentences were rated more highly than false sentences, and pragmatically "false" sentences were rated as intermediate. Furthermore, whether speaker identity also affected readers' evaluations differed across sentence types. For "some" sentences that were pragmatically underinformative, readers rated them better when they were led to believe that the sentences were written by a non-native speaker, but for the other three sentence types, identity did not play any role. Therefore, non-native speakers may enjoy a specific pragmatic advantage when others process the meaning of what they say.

This and other experiments examining the role of foreign speaker identity have used written text where participants were directly informed about the speaker's native or non-native status (cf., Fairchild & Papafragou, 2018). This leaves open the question of how speaker identity emerges in actual speech input; in real life conversation settings, listeners are rarely provided with explicit top-down information about speaker identity. To address this issue, in the present research we use spoken stimuli as a more direct test of speaker identity and its role in listeners' semantic and socio-pragmatic sentence processing. As a further contribution to this emerging literature, we examine how listeners perceive utterances by native and non-native speakers in real time.

We use two basic interpretative tasks that have been widely used both in research on scalar implicatures and in speech perception. In Experiment 1, we conduct a Sentence-Picture Verification task to see how listeners make speeded judgements about the truth-conditional content of an utterance (e.g., "Some of these circles are red") in the context of a concrete visual scene as an index of semantic processing. In Experiment 2, we use a Sentence-Only Verification task to examine listeners' judgements of common knowledge sentences (e.g., "All women are doctors") as a more open-ended measure of meaning. As in Fairchild and Papafragou (2018) and many other studies (e.g., Noveck & Posada, 2003; Bott & Noveck, 2004; De Neys & Schaeken, 2007), we present listeners in both experiments with four types of sentences: (1) true sentences with "all" (All-True), (2) false sentences with "all" (All-False), (3) true "some" sentences that carry a false pragmatic implicature (Some-False), and (4) true and pragmatically felicitous "some" sentences (Some-True). These sentences are presented in various degrees of accentedness; listeners hear sentences produced in either a native, a light foreign, or a heavy foreign accent.

Since foreign accents incur extra perceptual demands on listeners (e.g., Bent & Bradlow, 2003; Munro & Derwing, 1995), we predict that listeners' reaction times in both experiments will be slower for all sentence types when they are produced with a foreign accent. Of interest is whether listeners' semantic and socio-pragmatic interpretation of these sentences may differ depending on the native vs. non-native speaker identity. A first specific question is whether listeners will show lenience for falsehoods or infelicities in non-native speech. Of particular interest are the Some-False sentences, which, as mentioned earlier, are semantically true but pragmatically infelicitous as a result of a false scalar implicature. In previous studies, listeners have been found to differ in whether they interpret these sentences as literally true or pragmatically false (e.g., Bott & Noveck, 2004): could native vs. non-native speaker identity modulate this effect?

A second question is whether listeners' interpretation of native vs. non-native sentences might depend on the specific task, with greater variability when interpreting real-world sentences that invite richer socio-pragmatic reasoning

(Experiment 2) compared to the more straightforward demands of verifying a sentence against a simple visual scene (Experiment 1).

Experiment 1

Experiment 1 used a Sentence-Picture Verification task to examine potential effects of accents on listeners' speeded truth-value judgements of the propositional content of an utterance against a simple co-present geometric scene.

Participants

The final sample comprised 120 adult participants. All participants were monolingual native speakers of American English living in the United States of America and were recruited online via Prolific.

Materials and Procedure

The procedure for the Sentence-Picture Verification task was adapted from Zwaan, Stanfield, and Yaxley (2002). Participants first heard a quantified sentence with *all* or *some* (e.g., "All/Some of these circles are red") and then saw a picture (e.g., a picture with eight circles where five of the circles were red). While seeing the picture, participants had to press one of two buttons within a two-second time window to indicate whether the picture they were seeing matched the utterance the just heard (see Figure 1). Participants were told to respond as quickly and as accurately as possible. Participants had the option to press on their keyboard either "F" for "Match" or "J" for "Does Not Match".

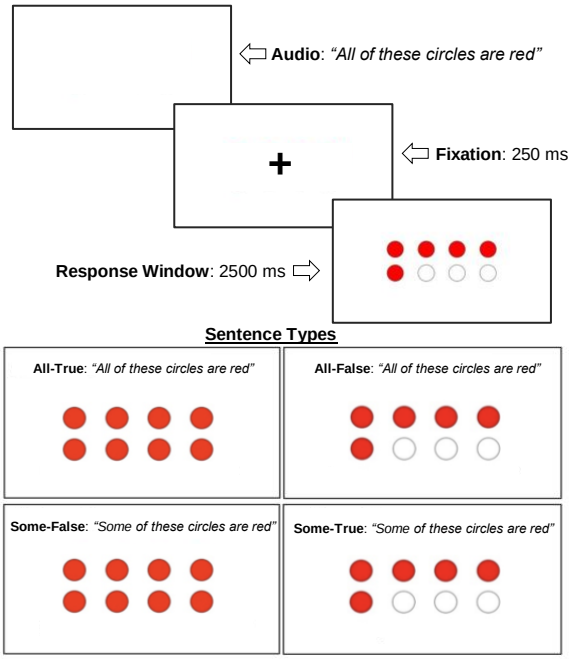

Figure 1: Sentence-Picture Verification task procedures and the 4 Sentence Types.

Sentence Type was manipulated within participants (i.e., All-True, All-False, Some-False, Some-True; note that the Some-False sentences were semantically true but conveyed a false implicature, "Some but not all"). Accent of the spoken utterances (Native, Light Foreign, Heavy Foreign) was manipulated between participants. The Native utterances were produced by a native female speaker of American English. Both the Light and Heavy Foreign accents were produced with a Chinese accent by a female speaker who was a native speaker of Mandarin Chinese. A norming study confirmed that all three accents differed significantly from each other in terms of their intelligibility: when asked "How easy is it to understand the speaker's English?", the Native accent was easier to understand than both the Heavy Foreign, $t(69) = 40.40$, $p < .001$, and the Light Foreign accent, $t(67) = 13.64$, $p < .001$, and the Light Foreign accent was easier than the Heavy Foreign accent, $t(143) = 6.07$, $p < .001$. A further norming study involving a separate sample of participants showed that, when asked "How close to a native speaker of English does the speaker sound?", the Native accent was more highly rated than both the Light, $t(67) = 19.52$, $p < .001$, and the Heavy Foreign accents, $t(69) = 46.90$, $p < .001$; however, the Light and Heavy Foreign accents did not differ in nativeness ratings, $t(143) = 1.45$, $p = .148$.

There was a total of 72 trials, with 18 trials for each Sentence Type. The trial sequence was pseudo-randomized. The pictures combined eight different shapes (e.g., circles, triangles, hearts) and nine different colors or patterns (e.g., red, blue, striped). All of the different shape and color-pattern combinations were counterbalanced. In conformance with general practice in research in speech perception, we also counterbalanced the trial order by including a list with the reversed order of the original trial order.

Results and Discussion

Generalized Linear Mixed-Effects Regression (GLMER) models were constructed for each Sentence Type to examine the effect of Accent on participants' accuracy rates. The baseline model included random slopes of sentence item by the effect of Accent; all other random slopes were excluded due to failure of the model to converge. Predictors (e.g., Accent) were added in a stepwise fashion to determine model fit; predictors that did not yield significant improvement were dropped before additional predictors were added. Based on our GLMER analyses, there were no significant main effects of Accent on listeners' accuracy rates for All-True, All-False, and Some-True utterances (see Table 1). However, a different pattern was observed for the Some-False utterances, where adding Accent as a predictor significantly improved model fit, χ^2 (1) = 36.88, $p < .001$. Here, the Some-False utterances produced with a Heavy Foreign accent were more likely to elicit pragmatic ("Do Not Match") responses compared to the same utterances produced with a Native accent.

Table 1: Percentage of correct (or pragmatically felicitous Some-False) responses in Experiment 1.

	All-True	All-False	Some-False (Pragmatic)	Some-True
Native	98.80%	96.40%	51.50%	95.80%
Light	97.10%	93.60%	53.90%	95.00%
Heavy	95.60%	93.80%	67.20%	92.90%

Further one-way ANOVAs were conducted for each Sentence Type to examine the effect of Accent on participants' response times (RTs) for correct responses (see Figure 2). Levene's adjusted *p*-values were used in cases of violation of sphericity, and the significance threshold (α= .05) for follow-up t-tests was Bonferroni-adjusted.

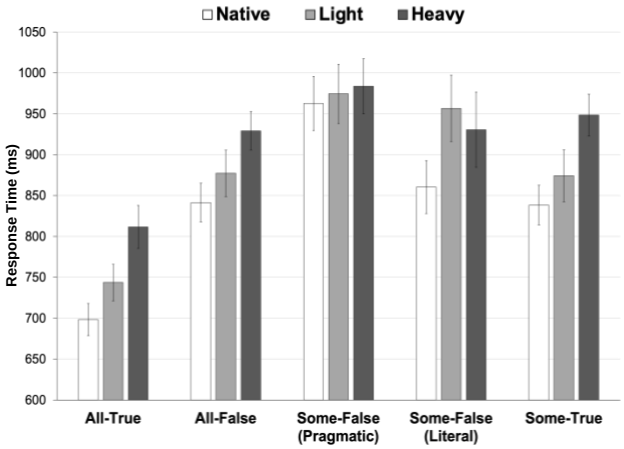

Figure 2: Response Times (in ms) for correct responses as a function of Accent and Sentence Type. Note that Some-False responses are for pragmatic responses.

For All-True sentences, there was a significant main effect of Accent, $F(2, 116) = 6.21$, $p = .003$. Follow-up t-tests found that the Native accent elicited faster RTs ($M = 698.35$, $SD = 124.50$) than the Heavy Foreign accent ($M = 811.63$, $SD = 166.45$), $t(78) = -3.45$, $p = .001$, but not the Light Foreign accent ($M = 743.55$, $SD = 140.09$), $t(77) = -1.52$, $p = .133$. However, the Light Foreign accent elicited marginally faster RTs than the Heavy Foreign accent, $t(77) = -1.96$, $p = .053$.

For All-False sentences, analyses revealed a marginally significant main effect of Accent, $F(2, 116) = 3.06$, $p = .051$, where the Native accent elicited faster RTs ($M = 841.36$, $SD = 150.58$) than the Heavy Foreign accent ($M = 929.17$, $SD = 149.85$), $t(78) = -2.61$, $p = .011$, but not faster than the Light accent ($M = 877.16$, $SD = 177.21$), $t(77) = -.97$, $p = .336$. There was also no RT difference between the Heavy vs. Light Foreign accents, $t(77) = -1.41$, $p = .163$

For the Some-False sentences, we analyzed listeners' RTs for both their responses that were pragmatically matching (i.e., "Do Not Match") and literally matching (i.e., "Match"). There were no main effects of Accent for either the pragmatic responses, $F(2, 86) = .10$, $p = .909$, or the literal responses, $F(2, 88) = 2.02$, $p = .138$.

For Some-True sentences, analyses also found a significant effect of Accent, $F(2, 117) = 4.20$, $p = .017$. Again, the Native accent elicited faster RTs ($M = 838.25$, $SD = 153.08$) than the Heavy accent ($M = 948.37$, $SD = 162.18$) $t(78) = -3.12$, $p = .003$, but not the Light accent ($M = 874.14$, $SD = 200.89$), $t(78) = -.90$, $p = .372$. RTs for the Light accent were also marginally faster than those in the Heavy Foreign accent, $t(78) = -.82$, $p = .073$.

Overall, the results indicate that listeners' Sentence-to-Picture (i.e., truth-value) judgements were largely unaffected by accent in terms of accuracy, even though the computation of judgments took longer in the presence of heavily accented speech. The one exception were Some-False sentences, where accent type did not affect RTs, and heavily accented utterances elicited more pragmatically matching responses than natively accented utterances, suggesting that listeners were less likely to view these sentences as literally true.

Experiment 2

Experiment 2 used a Sentence-Only Verification task to examine how listeners interpret sentence meanings in the presence of different speaker accents using sentences that were derived from everyday common knowledge (and were not accompanied by visual stimuli). This paradigm has been used in many past studies of pragmatic processing and scalar implicature (e.g., Noveck & Posada, 2003; Bott & Noveck, 2004; De Neys & Schaeken, 2007; Huang & Snedeker, 2009; Hunt et al., 2013; Tomlinson et al., 2013). By extending this paradigm to foreign accented speech, we asked whether different sentence types are processed differently as a function of the speaker's native vs. foreign identity.

Participants

A new sample of 120 adult participants was recruited from Prolific. As with Experiment 1, all participants were monolingual native speakers of English.

Materials and Procedure

As with the Sentence-Picture Verification task, participants heard a quantified sentence involving *all* or *some*, this time without pictures. Participants were asked to judge whether they "Agreed" or "Disagreed" with what the speaker said within a two-second timeframe immediately after hearing it (i.e., Press F for "Agree", Press J for "Disagree"). As in Experiment 1, participants were asked to respond as quickly and as accurately as possible.

Sentence Type was again manipulated as a within-participant variable: (1) All-True (e.g., "All snow is cold and can melt"), (2) All-False (e.g., "All women are doctors who went to medical school"), (3) Some-False (e.g., "Some fire is hot to the touch"), and (4) Some-True (e.g., "Some people have dogs as pets"). Accent was manipulated between participants (Native, Light Foreign, Heavy Foreign); the same speakers as in Experiment 1 provided the spoken sentences. There was a total of 72 trials, with 18 trials for each Sentence Types arranged in a pseudo-randomized order.

Results and Discussion

As in Experiment 1, separate GLMER models for each Sentence Type were constructed with random slopes of sentence items by the effect of Accent; again, other random factors were not included because the models failed to converge. Adding Accent into the model for All-True Sentence Type did not significantly improve model fit. However, there were main effects of Accent for both All-False, $\chi^2(1) = 10.47$, $p = .001$, and Some-True, $\chi^2(1) = 7.70$, $p = .006$: in both cases, listeners were more accurate when

agreeing or disagreeing with a statement for foreign accented sentences than for native accented sentences (see Table 2). For Some-False, however, analyses did not reveal a main effect of Accent.

Table 2: Percentage of correct (or pragmatically felicitous for Some-False) responses in Experiment 2.

	All-True	All-False	Some-False (Pragmatic)	Some-True
Native	77.50%	70.70%	25.60%	74.60%
Light	79.70%	90.50%	34.10%	96.20%
Heavy	75.40%	88.20%	22.60%	89.70%

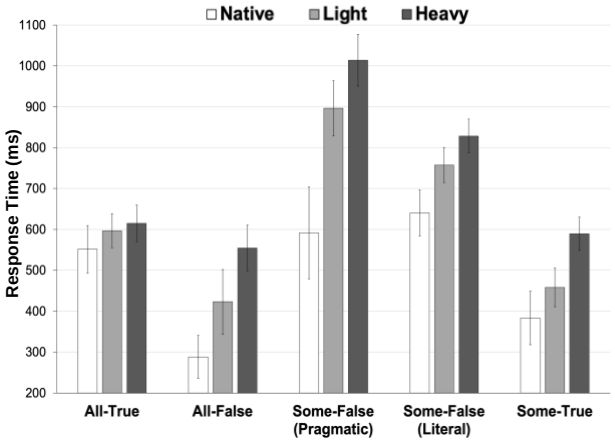

Figure 2: Response Times (in ms) for correct responses as a function of Accent and Sentence Type.

In terms of participants' correct judgement RTs (see Figure 3), analyses for All-True utterances did not find a significant main effect of Accent, $F(2, 107) = .44$, $p = .645$. For All-False sentences, however, there was a significant effect of Accent, $F(2, 97) = 3.72$, $p = .028$: utterances produced in the Native accent ($M = 288.08$, $SD = 281.49$) had faster RTs than utterances produced in the Heavy Foreign accent ($M = 554.47$, $SD = 320.26$), $t(60) = -3.46$, $p = .001$. However, there was no difference in RTs between the utterances with Native or Light Foreign accent ($M = 422.85$, $SD = 487.28$), $t(65) = -1.33$, $p = .189$, and between the utterances with Heavy and the Light Foreign accent, $t(69) = -1.32$, $p = .190$.

For Some-False utterances, analyses for pragmatic ("Disagree") responses revealed a significant main effect of Accent, $F(2, 79) = 6.78$, $p = .002$. Follow-up t-tests revealed that the Native accented sentences ($M = 591.32$, $SD = 526.70$) had faster RTs compared to both sentences produced in the Heavy Foreign accent ($M = 1013.76$, $SD = 335.10$), $t(48) = -3.45$, $p = .001$, as well as sentences produced in the Light Foreign accent ($M = 896.14$, $SD = 380.11$), $t(52) = -2.47$, $p = .017$. There was no difference in RT between the Heavy and Light Foreign accent, $t(58) = -1.26$, $p = .212$.

For the Some-False sentences eliciting literal ("Agree") responses, there was also a similar main effect of Accent, $F(2, 107) = 3.87$, $p = .024$. Here, Native accent ($M = 640.09$, $SD = 352.28$) elicited faster RTs than the Heavy Foreign accent ($M = 828.47$, $SD = 236.90$), $t(66.81) = -2.70$, $p = .009$, but not the Light Foreign accent ($M = 757.36$, $SD = 264.91$), $t(75) = -1.65$, $p = .104$, and there was no RT difference between the Light and Heavy Foreign accent, $t(69) = -1.18$, $p = .240$.

Finally, for Some-True sentences, there was also a significant main effect of Accent, $F(2, 102) = 3.86$, $p = .024$. Here again, utterances produced in the Native accent ($M = 383.18$, $SD = 382.23$) had faster RTs than the utterances produced in the Heavy accent ($M = 589.48$, $SD = 231.70$), $t(65) = -2.66$, $p = .010$, but not faster than those produced in the Light Foreign accent ($M = 458.21$, $SD = 290.06$), $t(70) = -.94$, $p = .348$, even though the utterances in the Light Foreign accent elicited marginally faster RTs than utterances in the Heavy Foreign accent.

As in Experiment 1, our RT results demonstrate that foreign accents incurred a perceptual cost on listeners' sentence processing, in line with in a wealth of previous studies (e.g., Bent & Bradlow, 2003; Munro & Derwing, 1995; Schmid & Yeni-Komshian, 1999). In addition, and orthogonally to this finding, listeners were more accurate at judging All-False and Some-True sentences when these were produced in a foreign accent; unlike Experiment 1, there was no difference in the interpretations of the Some-False sentences across the different accent contexts. We return to the interpretation of these results in the section below.

General Discussion

Recent years have seen a growing number of findings on the role of native vs. foreign accented speaker identity across various domains of language processing: syntactic errors are more likely to be ignored when produced in foreign-accented speech (e.g., Gibson et al., 2017; Hanulíková et al., 2012), pragmatic under-informativeness is tolerated more when attributed to a foreign accented speaker (e.g., Fairchild et al., 2020), and accented speakers who fail to disclose relevant information are less likely to be viewed as dishonest (e.g., Ip & Papafragou, 2021). All these findings suggest that listeners have different expectations for native vs. foreign speakers, and incorporate these expectations to interpret the sentence content produced by different speakers. The present experiments set out to test the boundaries and scope of listeners' reliance to stable cues from speaker identity (i.e., foreign non-native vs. native speakers) to interpret, in real time, the underlying meaning of an utterance.

We demonstrate that perceiving unfamiliar foreign accented speech carries perceptual costs (e.g., Bent & Bradlow, 2003; Lane, 1963; Munro & Derwing, 1995; Schmid & Yeni-Komshian, 1999): in both experiments, listeners were overwhelmingly slower at processing heavily accented sentences compared to those produced by a native speaker (see also Witteman, Weber, & McQueen, 2013). However, this cost did not negatively affect listeners' overall verification accuracy. In Experiment 1 where various outside parameters (e.g., world knowledge) were constrained through our manipulation of a visual context, listeners were more likely to correctly reject pragmatically infelicitous Some-False sentences when these were produced with a foreign accent (e.g., "Some of these circles are red" when all circles were red). Even in Experiment 2 that used more open-ended

everyday sentences (e.g., "All women are doctors who went to medical school"), we found no decrease in accuracy for accented speech, and in some cases (i.e., All-False and Some-True sentences), listeners were more accurate at judging the truth-value of the accented sentences. Therefore, perceptual costs associated with perception of accented speech may not necessarily permeate through listeners' processing accuracy, and in some contexts, the presence of foreign accents may even facilitate better processing of both the semantic and socio-pragmatic meaning of spoken sentences.

It is still an open question why sentence interpretation accuracy selectively improved in the presence of foreign accents despite the slowing down in response time. One explanation could be that participants deliberately adjusted their behavior when listening to accented utterances. Participants who heard sentences produced with an unfamiliar foreign accent might have slowed down not only because the foreign accent was harder to process, but also because they wanted or needed to be more careful, thus becoming more accurate. However, this explanation is not complete because it cannot account for the various nuances that we found in participants' response times and accuracy rates. Specifically, in Experiment 1 where only Some-False sentences showed more pragmatically matching responses in the non-native contexts, listeners showed no response time differences like they did with the other sentence types. Similarly, in Experiment 2, where accented utterances incurred slower response times for Some-False sentences, there were no significant improvements in listeners' pragmatic accuracy (although there is trend leaning towards more pragmatic interpretations for the lightly accented speaker). At the same time, previous research has shown that foreign accents cause listeners to process sentences in a less detailed way (Lev-Ari & Keysar, 2010), which further puts this explanation into question.

It is also unclear how listeners integrate speaker identity cues to choose between pragmatic vs. literal meanings of scalar implicture sentences (i.e., Some-False sentences). As already mentioned, when listeners were asked to make simple truth-value judgements during Sentence-Picture Verification task, they were significantly more likely to process these sentences with an upper-bounded pragmatic interpretation in the presence of heavily accented speech. However, this effect disappeared when listeners heard the everyday sentences in the Sentence-Only Verification task. It is therefore too early to conclude whether foreign accented speakers, in general, incur more pragmatic or literal interpretations, and whether this differs as a function of different experimental paradigms. Further research into these issues is warranted.

Overall, our findings have broader implications for accounts of how speaker identity affects non-native speech processing and social cognition. First, our experiments provide insights into how listeners process the pragmatics of native and non-native utterances. We show that accents are not only linguistic stimuli that affect speech processing, but also a rich set of social cues to speaker group identity that listeners must incorporate to make inferences about the speakers' communicative intentions (see Wilson & Sperber, 2004).

To conclude, we are far from understanding how listeners interpret the communicative intentions of different speakers, such as whether some types of speakers encourage more literal meanings, while others more pragmatic ones. Nevertheless, studying listeners' integration of native vs. non-native speaker identity cues, in real time, could provide insights into these issues.

Acknowledgments

We are extremely grateful to all the members of the Language and Cognition lab for their helpful comments. We would like to thank Ariel Mathis for her technical assistance, and Sarah Fairchild and Yue Ji for their help in the stimuli recordings.

References

Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *The Journal of the Acoustical Society of America, 114*(3), 1600-1610.

Bergen, L., & Grodner, D. J. (2012). Speaker knowledge influences the comprehension of pragmatic inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(5), 1450-1460.

Bott, L., & Noveck, I. A. (2004). Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language, 51*(3), 437-457.

Breheny, R. (2019). Scalar implicatures. In C. Cummins & N. Katsos (Eds.), *Oxford handbook of experimental semantics and pragmatics* (pp. 39-61). Oxford handbooks online.

De Neys, W., & Schaeken, W. (2007). When people are more logical under cognitive load: Dual task impact on scalar implicature. *Experimental Psychology, 54*(2), 128-133.

Fairchild, S., Mathis, A., & Papafragou, A. (2020). Pragmatics and social meaning: Understanding under-informativeness in native and non-native speakers. *Cognition, 200*, 104171.

Fairchild, S., & Papafragou, A. (2018). Sins of omission are more likely to be forgiven in non-native speakers. *Cognition, 181*, 80-92.

Gibson, E., Tan, C., Futrell, R., Mahowald, K., Konieczny, L., Hemforth, B., et al. (2017). Don't underestimate the benefits of being misunderstood. *Psychological Science, 28*(6), 703-712.

Gorbachev, M. S. (1987). *Perestroika: New thinking for our country and the world*. New York, NJ: Harper and Row Publishers.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J.J. Morgan (Eds.), *Syntax and semantics*. New York: Academic Press.

Grodner, D. J., Klein, N. M., Carbary, K. M., & Tanenhaus, M. K. (2010a). "Some," and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. *Cognition, 116*(1), 42-55.

Grodner, D. J., Klein, N. M., Carbary, K. M., & Tanenhaus, M. K. (2010b). "Some," and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. *Cognition, 116*(1), 42-55.

Hanulíková, A., van Alphen, P. M., van Goch, M. M., & Weber, A. (2012). When one person's mistake is another's standard usage: The effect of foreign accent on syntactic processing. *Journal of Cognitive Neuroscience, 24*(4), 878-887.

Huang, Y. T., & Snedeker, J. (2009). Semantic meaning and pragmatic interpretation in 5-year-olds: Evidence from real-time spoken language comprehension. *Developmental Psychology, 45*(6), 1723-1739.

Ip, M. H. K., & Papafragou, A. (2021). Listeners evaluate native and non-native speakers differently (but not in the way you think). In T. Fitch, C. Lamm, H. Leder, & K. Teßmar-Raible (Eds.), *Proceedings of the 43rd Annual Meeting of the Cognitive Science Society Conference 2021* (pp. 763-769).

Lane, H. (1963). Foreign accent and speech distortion. *The Journal of the Acoustical Society of America, 35*(4), 451-453.

Lev-Ari, S., & Keysar, B. (2010). Why don't we believe non-native speakers? the influence of accent on credibility. *Journal of Experimental Social Psychology, 46*(6), 1093-1096.

Lev-Ari, S., & Keysar, B. (2012). Less-detailed representation of non-native language: Why non-native speakers' stories seem more vague. *Discourse Processes, 49*(7), 523-538.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 45*(1), 73-97.

Noveck, I. A., & Posada, A. (2003). Characterizing the time course of an implicature: An evoked potentials study. *Brain and Language, 85*(2), 203-210.

Schmid, P. M., & Yeni-Komshian, G. H. (1994). The effects of accent and target predictability on perception of mispronunciations. *The Journal of the Acoustical Society of America, 95*(5), 2974-2974.

Skordos, D., & Papafragou, A. (2016). Children's derivation of scalar implicatures: Alternatives and relevance. *Cognition, 153*, 6-18.

Tomlinson, J. M., Bailey, T. M., & Bott, L. (2013). Possibly all of that and then some: Scalar implicatures are understood in two steps. *Journal of Memory and Language, 69*(1), 18-35.

Wilson, D., & Sperber, D. (2004). Relevance theory. In L. R. Horn, & G. Ward (Eds.), *The handbook of pragmatics* (pp. 607-632). Blackwell Publishing.

Witteman, M. J., Weber, A., & McQueen, J. M. (2013). Foreign accent strength and listener familiarity with an accent codetermine speed of perceptual adaptation. *Attention, Perception, & Psychophysics, 75*(3), 537-556.

Yildirim, I., Degen, J., Tanenhaus, M. K., & Jaeger, T. F. (2016). Talker-specificity and adaptation in quantifier interpretation. *Journal of Memory and Language, 87*, 128-143.

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science, 13*(2), 168-171.