Title
Is It All or Nothing? The Other Accent Effect in Talker Recognition

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

Authors
Yu, Madeleine
Johnson, Elizabeth

Publication Date
2023

Peer reviewed

Is It All or Nothing? The Other Accent Effect in Talker Recognition

Madeleine E. Yu (madeleine.yu@mail.utoronto.ca)
Department of Psychology, University of Toronto
Mississauga, ON, L5L 1C6, Canada

Elizabeth K. Johnson (elizabeth.johnson@utoronto.ca)
Department of Psychology, University of Toronto
Mississauga, ON, L5L 1C6, Canada

Abstract

Listeners often have trouble identifying other-accented talkers. Some suggest this Other Accent Effect (OAE) occurs only for non-native accents (e.g., Canadian English listeners experience it for Mandarin-accented English, but not Australian English). But the line between native and non-native accents can be difficult to distinguish, and past studies have confounded accent strength with accent type. Thus, we hypothesize that accent strength modulates the OAE. We predict a heavy non-native accent will elicit an OAE, whereas a light one will not. To test this, we presented native Canadian English listeners with voice line-ups of native Canadian English accented, non-native heavy Mandarin-accented, and non-native light Mandarin-accented talkers. Unsurprisingly, listeners performed better with Canadian English talkers than Mandarin-accented talkers. Crucially, listeners performed equally poorly with both heavy and light Mandarin-accented talkers. Thus, we found no evidence for our hypothesis; instead, we observe that even a weak non-native accent can elicit a strong OAE.

Keywords: talker recognition; Other Accent Effect; accented speech processing; speech perception; forensic earwitness

Introduction

Listeners sometimes find it very difficult to correctly identify talkers who speak in an accent different from their own (e.g., Stevenage et al., 2012, Johnson et al., 2018; Yu et al., 2021). This Other Accent Effect (OAE), however, is neither consistently observed nor well-understood. Why does the OAE occur in some studies but not others? Could it be that the presence of the effect is related to the differences between the accents that listeners are tested with? Here, we investigate the possibility that the OAE is more likely to occur for talkers who possess heavy accents.

Linguistic (i.e., what is being said) and indexical (i.e., who is speaking) information is processed in a simultaneous and integrated manner. For example, it is well known that familiarity with the language someone speaks affects how well they are later recognized (Thompson, 1987; Levi, 2019), and so it is unsurprising that some studies have suggested that familiarity with accents may similarly influence talker recognition ability. However, the mechanisms by which different accents influence the presence of the OAE still remain rather unclear as few studies have systematically investigated the effect.

One possible explanation for why the OAE is not always observed is that the effect is dependent on accent type. The particular accents (and particular listeners) used to test the effect differ across studies, making it difficult to generalize findings beyond a specific study. Some studies have used regional accents, whereas others have used non-native accents. While regional accents are learned from birth and reflect the manner of speaking of a particular region or community, non-native accents result from the porting of elements of a native language into a second language and may therefore differ substantially in underlying phonological structure (e.g., Bent, 2021; Cristia et al., 2012). Regional and non-native accents also differ in the types of social biases that they elicit (Adank et al., 2013) and in their perceived intelligibility (Adank et al., 2009; Yu et al., 2022). Any one of these differences can impact the presence of the OAE and indeed, recent investigations of the OAE have argued that the effect is observed with non-native accents, but not with regional accents (Yu et al., 2021; 2022). However, these two types of accents are also confounded with perceived accent strength, which, in fact, is perhaps a more reliable predictor of these differences than arbitrarily categorized accents are (e.g., see Cristia et al., 2012, for discussion). For instance, Spanish listeners find Standard Southern British English substantially more intelligible than Glaswegian English, despite the fact that both are categorized as regional variants of the same language (Stringer & Iverson, 2019). Also, it is observed that listeners are more likely to mistake difficult to comprehend regional accents as non-native accents (e.g., Ikeno & Hansen, 2006; 2007). Evidently, differences in accent strength, but not always in accent type, serve as strong markers for how accented speech is processed and perceived by listeners.

Research supports that linguistic processing is directly influenced by differences in accent strength even within the same accent. Listeners' word recognition is observed to become incrementally slower and less accurate as the perceived strength of accent increases (Porretta et al., 2016; Porretta & Kyröläinen, 2019). Thus, given the integrated and simultaneous nature with which indexical and linguistic information is processed, it might follow that differences in talkers' perceived accent strength also affect the processing of talker-related information. In other words, perhaps talkers who are perceived to possess heavy accents are also more difficult to later recognize compared to talkers who are perceived to possess light accents.

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society*.

In fact, there exists some support that differences in accent strength do affect talker recognition. In a talker identification training task featuring highly intelligible talkers, native speakers of American English demonstrated worse performance when identifying non-native accented talkers with a heavier Mandarin accent compared to their identification of non-native accented talkers with a lighter Mandarin accent (McLaughlin et al., 2019). These results suggest that listeners' perceptions of talker accent strength may serve as an important facilitator in accented talker recognition, though it should be noted while they did exclude participants with prior exposure to Mandarin, they did not indicate any exclusion criteria specific to exposure to Mandarin-accented English. Speech stimuli also did not differ between the training phase and test phase, making it difficult to determine if participants successfully generalized talker identity to speech stimuli when making identifications. Critically, the degree of talker accent strength was only observed to significantly affect accuracy when they did not control for the variation in identification of individual talkers, allowing for the possibility that their findings reflect both accent strength-related differences and individual talker-related differences in listeners' talker recognition.

The current study further examines the role of accent strength within the OAE by testing a considerably larger sample size (i.e., 72 in the current study compared to 24 in McLaughlin et al., 2019) on their ability to identify talkers who differ in perceived accent strength. We used a forensic-style voice line-up paradigm typically used in the talker recognition literature (see, e.g., Johnson et al., 2011; 2018; Yu et al., 2021; 2022) in which participants are familiarized with a talker's voice at the beginning of each trial, and then following a brief delay, they are asked to identify the same talker from a voice line-up that contains the target voice as well as three other distractor voices. In contrast to the talker identification training task paradigm, participants in this voice line-up paradigm received just a single brief exposure to each talker prior to test. Additionally, speech stimuli always differed between exposure and test ensuring that participants were more likely to generalize talker identity to new speech stimuli when making their identifications.

The task included voice line-ups for three accent types: native Canadian English, light Mandarin-accented English, and heavy Mandarin-accented English. We also took a measure of participants' confidence in their selections. To confirm that our heavy- and light-accented talkers perceptually differed in accent strength to listeners, we included a transcription task that featured an additional subset of speech stimuli produced by each of the talkers. We predicted that intelligibility (measured here as transcription accuracy) would differ between the accent types; in other words, listeners would experience greater difficulty understanding speech by the heavy-accented talkers than by the light-accented talkers.

In sum, we predicted that if differences in perceived accent strength are a major factor in the OAE, then we expected to observe a difference in how light- and heavy-accented talkers are recognized by listeners. Specifically, we predict that the OAE will be modulated by accent strength, with listeners experiencing the most difficulty recognizing heavy-accented Mandarin English talkers.

Method

Participants

To determine our sample size, we conducted a power analysis for a three-level mixed effect logistic regression in R. Based on the results of Yu et al. (2021; 2022), we expected native Canadian English-speaking listeners to better identify Canadian English talkers than Mandarin-accented English talkers. We also predicted gradient performance by perceived accent strength, so we set the conditions of the power analysis with participants demonstrating a 0.17 mean difference in the proportion of correct talker recognition between Canadian (0.72) and light Mandarin-accented English (0.55) talkers, and a 0.15 mean difference in the proportion of correct talker recognition between light and heavy Mandarin-accented English (0.40) talkers. The values for each condition were sampled from a normal distribution with a standard deviation of 0.20 mean by-participant proportion correct. The sample size was increased from 12 to 96 in steps of 12, with each step repeating 10,000 times. Eighty percent power was reached between 48 and 60 participants for both accent type comparisons. Thus, to ensure we had ample power to detect our predicted effect, we chose a sample size of 72.

A total of 72 native English-speaking adults were tested (M_{age} = 18.9 years; 56 female; 14 male; 2 non-binary). All participants learned English before the age of five. All participants reported using English at least 80% of the time and did not have routine exposure (in media or from a particular individual) to Mandarin-accented English. No hearing or vision impairments were reported at the time of testing. An additional 20 participants were excluded prior to the final analysis because they did not pass the pre-test practice trial (17), did not follow instructions (2), or experienced technical issues (1). This drop-out rate is typical for online experiments of this type.

Materials

Auditory stimuli consisted of recordings by four female adult talkers (M_{age} = 21.6 years) from each of the three accent groups. The selection process for Mandarin-accented talkers will be discussed in more detail further in this section. The script for sentence recordings was drawn from Johnson et al. (2011). Sentence length was controlled across accented speech (range of 15–21 syllables). All talkers were recorded reading the sentences in a neutral tone of voice in a double-walled, sound-attenuated Industrial Acoustics Company (IAC) booth using high-quality recording equipment (48kHz). All speech stimuli were normalized for root mean square amplitude in Praat 6.1.16 (Boersma & Weenink, 2020).

The Canadian English talkers were from the Greater Toronto Area and all learned English before the age of 5. The Mandarin-accented English talkers all learned English after the age of 5, and while not from a single metropolitan area, were confirmed by two native speakers of Mandarin to be native speakers of standard Mandarin. All had a highly perceptible Mandarin accent when speaking English.

Perceptions of non-native accent strength were obtained from ratings made by a separate sample of 15 Canadian English dominant adult listeners for 9 Mandarin-accented English talkers on a 6-point rating-scale (1 = subtle accent, 6 = heavy accent; see Yu et al., 2019 for more details on the accent rating task). For the purposes of the current experiment, light- and heavy-accented talker sets were constructed by assigning the four talkers with the lowest accent strength ratings (range = 1.79 – 3.45) to the "light Mandarin accent" group and the four talkers with the highest accent strength ratings (range = 3.85 – 4.66) to the "heavy Mandarin accent" group.

Prior studies involving forensic-style voice line-ups have closely matched talkers in their mean F0 (e.g., Johnson et al., 2011; 2018; Yu et al., 2021; 2022). To keep this study consistent with existing work and to aid the comparability of our findings, the mean fundamental frequency (F0) for each talker was shifted using the built-in pitch-shifting function in Praat to match the mean F0 of the respective talker set of each accent group. The recordings still sounded naturalistic following pitch shifts. As in past studies of this type, there were no other obvious impressionistic differences (e.g., excessive breathiness) in voice quality across talkers in each set. The resynthesized mean F0 for the Canadian English talkers, light Mandarin-accented talkers, and heavy Mandarin-accented talkers were 203.5 Hz, 202.1 Hz, and 204.1 Hz respectively. The talker sets we used did not differ in relative variability of mean F0 and duration (following prior work of Johnson et al., 2011; 2018; Yu et al., 2021; 2022, F-values were calculated as the ratio of the two variances, and comparisons for these two acoustic features were below the critical F-value of 9.28).

Table 1: Relative Acoustic Variability of Talker Sets in F-Values

Accent Pair	Duration	Mean F0
CAN-HeavyMand	1.04	1.26
CAN-LightMand	1.26	1.22
HeavyMand-LightMand	1.22	1.53

To validate our classification of talkers as having a heavy or light accent, we collected intelligibility ratings. Since heavier accents are typically less intelligible than lighter accents, we anticipated that the talkers we classified as heavily accented would be less intelligible than the talkers we classified as lightly accented (McLaughlin, 2019; Porretta & Kyröläinen, 2019). After completing the main voice identification part of the study, each participant was presented with recordings of 8 different sentences from each accent type embedded in noise (0 SNR) for a total of 24 unique sentences in randomized order. None of these sentences were used in the voice identification task. Each of the 12 talkers produced two different sentences, so that each participant heard each talker twice. Across all participants, recordings were presented in each accent condition an even number of times.

Transcriptions by all participants were manually coded for errors by a single coder blind to the conditions of the task. A transcription error involved any phonologically distinct word that deviated from the original content of the target sentence. Homophonous variants (e.g., 'son' instead of 'sun'), or transcriptions that deviated from the target word by no more than one letter and could not be considered a distinctly different word (e.g., 'deligation' instead of 'delegation'), were identified as misspellings and typos and were subsequently hand-corrected by the coder. As expected, mean transcription accuracy was lowest for heavy Mandarin-accented English compared to the other English variants ($M_{Canadian}$ = 0.83, SD = 0.21; $M_{Light\text{-}Mandarin}$ = 0.76, SD = 0.27; $M_{Heavy\text{-}Mandarin}$ = 0.68, SD = 0.28). Moreover, paired *t*-tests revealed that transcription accuracy was significantly worse for heavy Mandarin-accented speech than for light Mandarin-accented speech, $t(71) = 6.62$, $p < .001$, and Canadian, $t(71) = 13.14$, $p < .001$, speech. Mean transcription accuracy was also lower for light Mandarin-accented speech than it was for Canadian, $t(71) = 5.71$, $p < .001$, speech. These results indicate that our heavy- and light-accented talkers differ in intelligibility, and thus validate our classification of talkers as heavily or lightly accented.

Procedure

The task was created and hosted via the Gorilla Experiment Builder (www.gorilla.sc; Anwyl-Irvine et al., 2020). Prior to testing, participants were instructed to complete the task in a quiet room free of distractions and to use headphones. A headphone screening test based on an antiphase discrimination task (Woods et al., 2017) was administered to ensure that participants were wearing headphones before they were allowed to progress to the main task. All participants took about 1 hour to complete the task in its entirety.

Following the design of prior talker recognition tasks (e.g., Yu et al., 2021; 2022), participants were asked to identify talkers from a forensic-style voice identification task. All participants heard three types of accented voice line-ups: Canadian English, light Mandarin-accented English, and heavy Mandarin-accented English.

At the start of each the 12 trials, participants were familiarized with a talker who produced a pair of sentences (interstimulus interval [ISI] between sentences = 300ms) upon a participant-initiated mouse click. Following 500ms of silence, a one-minute distractor video clip which featured instrumental music and sound effects, but no speech (Fecher & Johnson, 2018; Yu et al., 2021; 2022) was displayed. Next, participants were presented with a four-voice line-up that contained the target talker and three distractor talkers in pseudo-randomized order. The distractor talkers were always from the same accent group as the target, and each produced a single sentence identical in content. Following a participant-initiated mouse click and 500ms of silence, the sentence recording by the left-most talker played and automatically advanced to the recording by the next talker to the right following a 1000ms ISI. After the presentation of all talkers, participants were asked to judge which of the four talkers was the target. Participants were told to guess if unsure. Upon making a talker selection, participants were given the option to change their selection before making a final confirmation. After each selection, participants reported their confidence with their decision using a continuous sliding scale between 0 (not confident at all) and 100 (highly confident). A single practice trial, featuring acoustically distinct voices exclusive from the rest of the task, preceded the test trials and ensured that participants understood the task. No feedback on correctness was provided to participants on any of the trials. A four-minute break occurred halfway through the task.

Participants heard 14 different sentences (2 familiarization sentences + 12 different line-up sentences) throughout the entirety of the voice line-up task; all talkers in all accent conditions produced the same two familiarization sentences. This was done to ease processing of the accented speech (see Baese-Berk et al., 2021; Bradlow & Alexander, 2007; Davis et al., 2005, for similar manipulations of comprehension load), and to increase the likelihood that participants successfully encode talker-related information during this phase.

Twelve unique sentences were featured across the line-ups. In each line-up the same sentence was produced by all four talkers. For all participants, the position of the target talker was fully counterbalanced such that it was equally likely for the target to appear in any of the three positions. Trial order was also pseudo-randomized so that no accent condition occurred more than twice-in-a-row. The order of accent conditions was fully counterbalance across all participants.

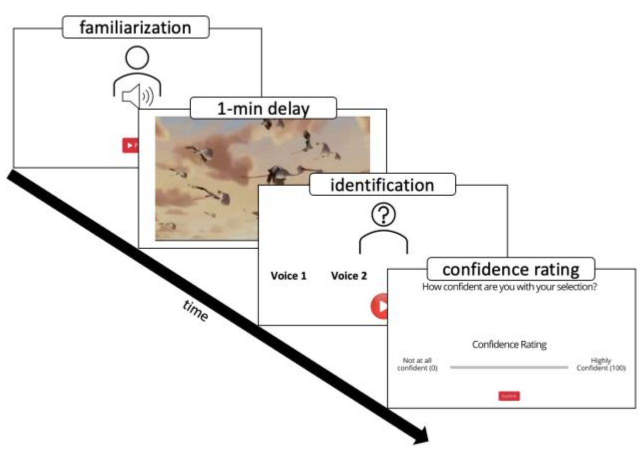

Figure 1: An example trial, in order of presentation

Results

Accuracy

Listeners' average talker recognition by accent type is shown in Figure 2.

To compare the effect of accent type on talker recognition accuracy we fit a generalized logistic mixed-effects regression model to the data using the *glmer* function in the *lme4* package Version 1.1-31 (Bates et al., 2015) in R. The model included the binary response variable, talker recognition accuracy (1 = correct response). Accent type was included as the independent variable and was helmert-coded to allow us to compare: (1) Canadian vs. light Mandarin and heavy Mandarin accents, and (2) light Mandarin vs. heavy Mandarin accent. The maximal random effects structure that would converge was implemented and included a random intercept for talker and a random by-participant intercept and slope for accent type.

Listeners in all accent conditions performed above chance (all *p*s < .001). There was also an effect of accent type on listener performance. Listeners were significantly more accurate at recognizing native Canadian English talkers compared to both Mandarin accent type conditions ($M_{accuracy}$ = 0.667 vs. 0.452), $\beta = 0.91$, $SE = 0.26$, $z = 3.51$, $p < .001$. There was no significant difference in accuracy between light Mandarin and heavy Mandarin accents ($M_{accuracy}$ = 0.469 vs. 0.434), $\beta = 0.14$, $SE = 0.30$, $z = .476$, $p = .634$. One-sample *t*-tests revealed a significant difference in talker recognition between Canadian and light Mandarin accents ($M_{accuracy}$ = 0.667 vs. 0.469), $t(71) = 5.50$, $p < .001$, as well as between Canadian and heavy Mandarin accents ($M_{accuracy}$ = 0.667 vs. 0.434), $t(71) = 5.82$, $p < .001$.

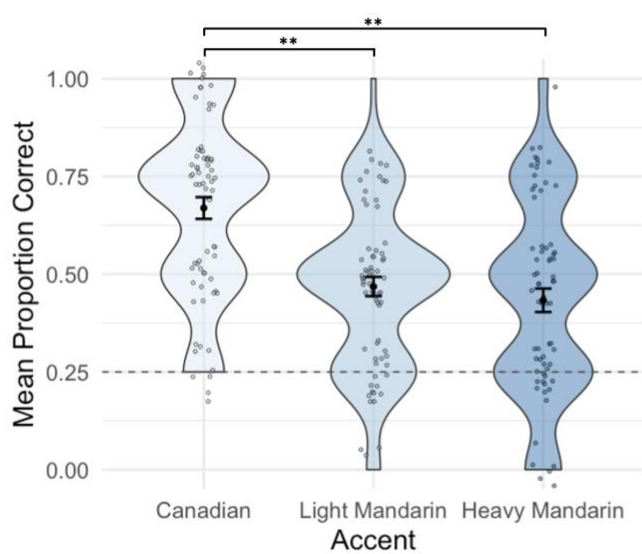

Figure 2: Mean proportion of correct talker recognition by accent type. Error bars indicate SE of by-participant means, and the dashed line indicates chance level performance.

Confidence

Listeners' confidence in talker recognition by accent type is shown in Figure 3.

To compare the effect of accent type on confidence, we fit a linear mixed-effects regression model to the data using the *lmer* function in the *lmerTest* package Version 3.1-3 (Kuznetsova et al., 2017) in R. The model included confidence rating as the continuous response variable and accent type as the independent variable. Accent type was forward difference-coded to allow for adjacent comparisons: (1) Canadian vs. light Mandarin accent, and (2) light Mandarin vs. heavy Mandarin accent. The maximal random effects structure that would converge was implemented and included a random intercept for talker and a random by-participant intercept and slope for accent type.

There was a significant difference in confidence for talker recognition between Canadian and light Mandarin accents ($M_{confidence}$ = .706 vs. .585), $\beta = 0.12$, $SE = 0.03$, $t = 4.64$, $p < .001$, and there was a significant difference in confidence between light and heavy Mandarin accents ($M_{confidence}$ = .585 vs. .529), $\beta = 0.06$, $SE = 0.02$, $t = 2.34$, $p < .05$. A one-sample *t*-test confirmed there was a significant difference in confidence between Canadian and heavy Mandarin accents ($M_{confidence}$ = .706 vs. .529), $t(71) = 9.50$, $p < .001$.

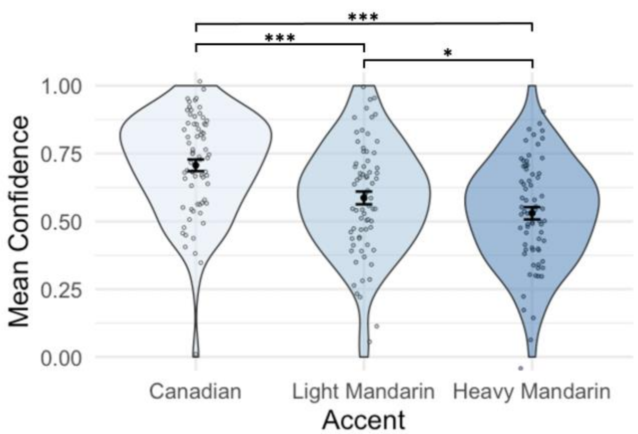

Figure 3: Mean confidence in talker recognition by accent type. Error bars indicate SE of by-participant means.

The observed difference in overall confidence, but not in performance, between light and heavy Mandarin-accented English suggests that the relationship between confidence and accuracy may differ by accent type. To investigate this relationship, we fit a generalized logistic mixed-effects model to our data using the *glmer* function in the *lme4* package in R. The model included talker recognition accuracy as the binary response variable, and accent type, confidence, and their interaction, were included as independent variables. Accent type was forward difference-coded as in the previous models. The maximal random effects structure that would converge was implemented and included a random intercept for talker and a random by-participant intercept and slope for accent type.

There was a main effect of confidence, $\beta = 2.27$, $SE = 0.39$, $z = 5.86$, $p < .001$, indicating that the correct talker was chosen more often when higher confidence was reported. However, there was no significant interaction between confidence and any of the accent type comparisons (all *p*s n.s.), indicating that in all accent conditions, participants selected the correct talker when they were more confident and the incorrect talker when they were less confident.

Discussion

Past work suggests that the Other Accent Effect (OAE) is not consistently observed with all "other" accents, even when identical paradigms are used to test performance (e.g., the OAE is present with Mandarin-accented English but not with Australian English; Yu et al., 2021; 2022). Why is the OAE observed in some cases and not in others? Here, we investigated whether the presence of the effect could be explained by differences in talkers' accent strength. To test this hypothesis, participants were presented with a forensic-style voice identification task that involved three different line-ups: Canadian English, light Mandarin-accented English, and heavy Mandarin-accented English talkers. The difference in accent strength between the two Mandarin-accented line-ups was identified via participants' confidence ratings and subsequently confirmed by their transcription performance. As predicted, listeners recognized Canadian English talkers significantly better than Mandarin-accented talkers, replicating prior work which finds the OAE is prominent with non-native accents (Yu et al., 2021; 2022). Contrary to our predictions, however, there was no difference in performance between the strong and weak non-native Mandarin accents. Overall, our findings suggests that while accent strength differences do affect perceptions of talker intelligibility and confidence in talker selections, we find no evidence that the OAE is more likely to be observed with heavy-accented talkers.

Given our current findings, how can we explain prior work, such as McLaughlin et al. (2019), which reports a decrement in talker recognition accuracy the more that an accent diverges from that of native talkers? As previously mentioned, the current study investigated the role of accent strength in the OAE in several different ways from prior work. First, in addition to a substantially larger sample size than that previously used, the current task also used a separate set of speech stimuli during the familiarization and test phase of the trial requiring participants to generalize voice information of the talkers when identifying them from each line-up. Secondly, when random effects for individual talker differences are omitted from our analyses, as they were in McLaughlin et al. (2019), all our effects remained consistent. And thirdly, in contrast to a talker identification training task, the brief nature with which listeners were familiarized to talkers in the current task is perhaps more analogous to a real-world situation where individuals receive limited exposure to the speech of a talker before they might be required to identify them.

It is certainly possible that current task demands influenced participant performance. Participants were not required to learn the voices of the talkers prior to test, so it is possible that attentional factors superseded the processing of accent strength differences, resulting in equally worse processing of all non-native talkers. Non-native accents, like Mandarin-accented English in this case, also tend to be more negatively perceived compared native accents (e.g., Lindemann, 2002; Baquiran & Nicholadis, 2020; Boduch-Grabka & Lev-Ari, 2021; Lev-Ari & Keysar, 2010; Oppenheimer, 2008) and this is argued to modulate attention-based encoding (Sumner & Kataoka, 2013). Specifically, participants may have allocated less attentional resources to encode voice information when they detected a non-native accent, regardless of the degree of accent strength. Some research indicates however, that light non-native accented talkers may be evaluated just as positively as native accented talkers (see Hendriks & van Meurs, 2021, for discussion). In-trial confidence ratings also suggest that listeners in the current study were sensitive to accent strength differences between the two Mandarin-accented talker sets.

Importantly, although we found no evidence that accent strength modulates the presence of the OAE in talker recognition performance, we did find that accent strength differences mattered in some instances. For example, listener confidence in talker selections were sensitive to differences in accent strength even though performance was not. This disconnect between overall confidence and performance across conditions is in line with the findings of Yu et al. (2021) and supports that the strength of listener confidence depends on both accent type and strength. This observation is important, particularly for forensic contexts, as it demonstrates that earwitnesses' confidence in their testimony might not always align with actual identification accuracy, especially for non-native talkers with light accents. Listeners' transcription accuracy was also sensitive to differences in accent strength, providing further support for the idea that accent strength directly affects linguistic processing (Porretta et al., 2016; Porretta & Kyröläinen, 2019). Interestingly, the same listeners' processing of talker identity was not influenced by accent strength in the same way, which suggests that alternate factors (e.g., social biases, attention, processing load, task demands) may have affected listener performance during the talker recognition task. Future investigation should explore whether increasing participant motivation to attend to all talkers equally would facilitate accent strength-related differences in recognition.

In sum, this work furthers our understanding of the mechanisms that underly the OAE, as well as talker recognition in general. Our findings support existing work demonstrating that the OAE typically occurs with non-native accents, such as Mandarin-accented English. However, we found no evidence that accent strength modulates the OAE, which suggests that other variables drive the presence of the effect. Further work is needed to disentangle how external factors, such as listeners' motivation and social biases, in addition to task demands, might ultimately determine the presence of the OAE in talker recognition.

Acknowledgments

We would like to thank Alexandra Hawse and the Child Language and Speech Studies Lab for their support. This work was supported by grants from the Social Sciences and Humanities Research Council and the Natural Sciences and Engineering Research Council.

References

Adank, P., Evans, B. G., Stuart-Smith, J., & Scott, S. K. (2009). Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *Journal of Experimental Psychology: Human Perception and Performance, 35*(2), 520.

Adank, P., Stewart, A. J., Connell, L., & Wood, J. (2013). *Accent imitation positively affects language attitudes. Frontiers in psychology, 4,* 1–10.

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods, 52*, 388–407. https://doi.org/10.3758/s13428-019-01237-x

Baese-Berk, M. M., Bent, T., & Walker, K. (2021). Semantic predictability and adaptation to nonnative speech. *JASA Express Letters, 1*(1), 015207.

Baquiran, C., & Nicoladis, E. (2020). A doctor's foreign accent affects perceptions of competence. *Health Communication, 35*(6), 726–730. https://doi.org/10.1080/10410236.2019.1584779

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

Baumann, O., & Belin, P. (2010). Perceptual scaling of voice identity: Common dimensions for different vowels and speakers. *Psychology Research, 74,* 110– 120.

Bent, T., Holt, R. F., Van Engen, K. J., Jamsek, I. A., Arzbecker, L. J., Liang, L., & Brown, E. (2021). How pronunciation distance impacts word recognition in children and adults. *The Journal of the Acoustical Society of America, 150*(6), 4103-4117.

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

Boersma, P. and Weenink, D. (2020). "Praat: doing phonetics by computer" [Computer program], version 6.1.16, http://www.praat.org/.

Bradlow, A. R., & Alexander, J. A. (2007). Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners. *The Journal of the Acoustical Society of America, 121*(4), 2339–2349. https://doi.org/10.1121/1.2642103

Cristia, A., Seidl, A., Vaughn, C., Schmale, R., Bradlow, A., & Floccia, C. (2012). *Linguistic processing of accented speech across the lifespan. Frontiers in psychology, 3*, 479.

Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A., Taylor, K., & McGettigan, C. (2005). Lexical information drives perceptual learning of distorted speech: evidence from the comprehension of noise-vocoded sentences. *Journal of Experimental Psychology: General, 134*(2), 222.

Fecher, N. & Johnson, E. K. (2018). Effects of language experience and task demands on talker recognition by children and adults. *The Journal of the Acoustical Society* *of America, 143*(4), 2409–2418. https://doi.org/10.1121/1.5032199

Hendriks, B., van Meurs, F. (2021). Foreign Accent Strength in English. In: Mohebbi, H., Coombe, C. (eds) Research Questions in Language Education and Applied Linguistics. Springer Texts in Education. Springer, Cham. https://doi.org/10.1007/978-3-030-79143-8_82

Ikeno, A., & Hansen, J. H. (2007). The effect of listener accent background on accent perception and comprehension. *EURASIP Journal on Audio, Speech, and Music Processing*, *2007*, 1-8.

Ikeno, A., & Hansen, J. H. (2006). Perceptual recognition cues in native English accent variation: listener accent, perceived accent, and comprehension. *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings, 1*, 401-404.

Johnson, E. K., Bruggeman, L., & Cutler, A. (2018). Abstraction and the (misnamed) language familiarity effect. *Cognitive Science, 42*(2), 633–645. https://doi.org/10.1111/cogs.12520

Johnson, E. K., Westrek, E., Nazzi, T., & Cutler, A. (2011). Infant ability to tell voices apart rests on language experience. *Developmental Science, 14*(5), 1002–1011. https://doi.org/10.1111/j.1467-7687.2011.01052.x

Kinzler, K. D., Shutts, K., DeJesus, J., & Spelke, E. S. (2009). Accent trumps race in guiding children's social preferences. *Social Cognition, 27*(4), 623–634. https://doi.org/10.1521/soco.2009.27.4.623

Kuznetsova A., Brockhoff P.B. and Christensen R.H.B. (2017). "lmerTest Package: Tests in Linear Mixed Effects Models." *Journal of Statistical Software*, *82*(13), pp. 1–26. doi: 10.18637/jss.v082.i13.

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

Levi, S. V. (2019). Methodological considerations for interpreting the Language Familiarity Effect in talker processing. *WIREs Cognitive Science, 10*(2), e1483. https://doi.org/10.1002/wcs.1483

Lindemann, S. (2002). Listening with an attitude: A model of native-speaker comprehension of non-native speakers in the United States. *Language in Society, 31*(3), 419-441.

McLaughlin, D. E., Carter, Y. D., Cheng, C. C., & Perrachione, T. K. (2019). Hierarchical contributions of linguistic knowledge to talker identification: Phonological versus lexical familiarity. *Attention, perception & psychophysics, 81*(4), 1088–1107. https://doi.org/10.3758/s13414-019-01778-5

Oppenheimer, D. M. (2008). The secret life of fluency. *Trends in Cognitive Science, 12*, 237–241.

Pérez-Ramón, R., Lecumberri, M. L. G., & Cooke, M. (2022). Foreign accent strength and intelligibility at the segmental level. *Speech Communication*, *137,* 70-76.

Perrachione, T. K., Dougherty, S. C., McLaughlin, D. E., & Lember, R.A. (2015). The effects of speech perception and speech comprehension on talker identification. *18th International Congress of Phonetic Sciences.*

Porretta, V., & Kyröläinen, A. J. (2019). Influencing the time and space of lexical competition: The effect of gradient foreign accentedness. Journal of Experimental Psychology: *Learning, Memory, and Cognition, 45*(10), 1832.

Porretta, V., Tucker, B. V., & Järvikivi, J. (2016). The influence of gradient foreign accentedness and listener experience on word recognition. *Journal of Phonetics, 58*, 1-21.

Stevenage, S. V., Clarke, G., & McNeill, A. (2012). The 'other-accent' effect in voice recognition. *Journal of Cognitive Psychology, 24*(6), 647–653. https://doi.org/10.1080/20445911.2012.675321

Stringer, L., & Iverson, P. (2019). Accent intelligibility differences in noise across native and nonnative accents: Effects of talker–listener pairing at acoustic–phonetic and lexical levels. *Journal of Speech, Language, and Hearing Research, 62*(7), 2213-2226.

Sumner, M., & Kataoka, R. (2013). Effects of phonetically-cued talker variation on semantic encoding. *The Journal of the Acoustical Society of America, 134*(6), EL485-EL491.

Thompson, C. P. (1987). A language effect in voice identification. *Applied Cognitive Psychology, 1*(2), 121–131. https://psycnet.apa.org/doi/10.1002/acp.2350010205

Woods, K. J. P., Siegel, M. H., Traer, J., & McDermott, J. H. (2017). Headphone screening to facilitate web-based auditory experiments. Attention, Perception, and *Psychophysics, 79*(7), 2064–2072. https://doi.org/10.3758/s13414-017-1361-2

Yu, M., Schertz, J., & Johnson, E. (2022). Do I need to repeat myself? Getting to the root of the Other Accent Effect. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 44. Retrieved from https://escholarship.org/uc/item/2kr0g3fv

Yu, M.E., Schertz, J., & Johnson, E.K. (2021). The other accent effect in talker recognition: now you see it, now you don't. *Cognitive Science*, *45*(6). doi.org/10.1111/cogs.12986

Yu, M. (2019). Re-evaluating the other accent effect in talker recognition (Publication No. 27541135) [Master's Thesis, University of Toronto]. ProQuest Dissertations & Theses Global