The Malleability of Listener Judgments of Second Language Speech

Kym Taylor Reid, Pavel Trofimovich, and Mary Grantham O'Brien

Abstract In research and practice contexts, assessment of second language (L2) speakers often falls to listeners (e.g., naïve listeners, trained assessors, language teachers) who, as members of their respective linguistic communities, might be influenced by various social biases, both positive and negative. However, it is presently unclear whether listeners are immune to external social biases and how the potential impact of these biases on listener assessments of L2 speech can be mitigated. The goal of this chapter is therefore to review our recent empirical work examining the malleability of listeners' evaluations of L2 speech for several speech dimensions, including accentedness, comprehensibility, and fluency. We first review existing research regarding the impact of negative and positive social bias on naïve listeners' and language teachers' evaluations of L2 speech. We then discuss the roles of various interventions, such as task practice and rater training, as ways of mitigating social bias effects on listeners' speech assessments. We conclude by discussing possible implications of our findings for the teaching, research, and assessment of L2 speech.

Keywords Speech rating · Assessment · L2 English · Rating stability · Social bias · Task practice

K. Taylor Reid (✉) · P. Trofimovich
Concordia University, Montreal, Canada
e-mail: kym.taylor@concordia.ca

P. Trofimovich
e-mail: pavel.trofimovich@concordia.ca

M. G. O'Brien
University of Calgary, Calgary, Canada
e-mail: mgobrien@ucalgary.ca

1 Introduction

Social contact provides opportunities for people to form impressions about others. A person's speech, in particular, offers unique insight, providing listeners with various clues about the speaker's amiability, mood, place of origin, or education level (Yzerbyt et al., 1994). However, listeners' perceptions can also conflict with their expectations that originate from specific prior experiences or internalized beliefs. When such disparity occurs, listeners can either accept the new information or default to their initial beliefs, leading to an overgeneralization known as stereotyping or bias (Stroebe & Insko, 2013). This overgeneralization, based on unfounded beliefs and misinformation, can then be used as justification for marginalizing others. In our recent work, we have focused on socially constructed biases in the domain of second language (L2) speech to explore this societal problem. Our specific goal has been to determine the extent to which social biases occur in the context of listeners evaluating L2 speech, so that these biases can be reduced through various interventions, including perspective taking and task practice.

2 Literature Review

2.1 Origins of Social Bias

Humans naturally seek to identify themselves as members of various in-groups (i.e., groups to which they belong), often as a means of improving their chances of being perceived positively by others (Mullin & Hogg, 1999). This includes using speech (defined broadly as segmental and suprasegmental characteristics of speech contributing to listener perceptions of a speaker's accent) to categorize the speaker as being a member (or an outsider) of specific groups (Bourhis et al., 2012; Giles & Watson, 2013). In this way, a person's accent tells the story of who they are, where they have been (Matsuda, 1991) and, most importantly, where they belong. An example of this accent-based categorization can be found in Labov's (1972) classic study in which speakers showcased a unique identity as in-group members using their pronunciation (i.e., variation in vowel quality). More recent work has confirmed that people indeed use language to underscore between-group similarities and differences (and their associated stereotypes) and make judgments about others (Bourhis et al., 2012; Dragojevic et al., 2016; Ryan, 1983; Wigboldus et al., 2005).

Sometimes, speech-based distinctions are linked to positive listener attitudes. Speakers in Dalton-Puffer et al.'s (1997) study, for instance, who spoke British English with a Received Pronunciation accent, were rated by L2 learners as more courteous, educated, and organized than those with other native accents. Similarly, Heaton and Nygaard (2011) found that listeners from across the United States rated English speakers with an American southern accent to be more sociable than those with a standard American accent. When presented in isolation, these stereotypical

attitudes do not appear to be particularly harmful but, in reality, the upgrading of one group often leads to the downgrading of another. In Heaton and Nygaard's study, for instance, though speakers with a southern accent were found to be more sociable than those with a standard American accent, the roles were reversed for ratings of intelligence. So, depending on which factors of the speakers are being assessed, someone is almost always in the out-group.

Negative attitudes that follow from such judgments can be particularly damaging when expressed by majority groups against minority status speakers, especially considering that humans exhibit a natural predisposition to "dislike the unlike" (Kagedan, 2020, p. 5). The detrimental effect of speech-based biases is perhaps best illustrated in workplace studies such as Ryan et al. (1977), where native English listeners perceived heavily-accented Spanish speakers to be of lower occupational status and to be less friendly than their less-accented counterparts. In fact, job applicants from minority groups are about 50% less likely to be invited to a job interview (Zschirnt & Ruedin, 2016), and immigrants (most of whom speak with a discernible foreign accent) tend to be underemployed (Krahn et al., 2000), with foreign accents considered more suitable for low than for high prestige jobs (Brennan & Brennan, 1981).

Attitudes toward specific cultural groups, societal norms regarding minority speakers, and the role of language within a particular society—including how it is used in education, politics, and the media—can all influence listener perception and stigmatization of L2 speakers (Gluszek & Dovidio, 2010). For instance, in an investigation of consumer responses to radio advertising, Lalwani et al. (2005) found that residents of Singapore believed standard English speakers to be more credible spokespersons than their fellow speakers of Singlish (Singapore English), even when pitching Singaporean products. To further complicate the issue of perception, at least some of these attitudes and behaviors can be attributed to listener expectations of speech before it is even heard (Lindemann & Subtirelu, 2013). In Babel and Russell's (2015) study investigating the effects of face priming on ratings of speech samples produced by native speakers of Canadian English, speech samples primed with photos of self-identified Chinese Canadians were rated as less intelligible and more accented than those primed with photos of self-identified White Canadians.

In another example in which expectations have been shown to fuel speech attitudes, Lindemann (2002) paired native English speakers with native Korean speakers for an interactive task in English. Some native speakers who were found to have negative attitudes toward Koreans prior to the interaction neglected to acknowledge communication from their partners and, in some cases, even withheld vital information. These same speakers subsequently reported their interactions to be less successful than native speakers who held positive attitudes toward Koreans. In fact, the root of most communication problems encountered during the task involved the "negative attitude" participants, which highlights the significant role of attitudes in successful communication, regardless of the language ability or proficiency of interlocutors.

When linguistic differences are used to judge speakers based on imagined or preconceived ideas, biases can become even more problematic. This is known as

reverse linguistic stereotyping (Kang & Rubin, 2009), or the process by which general attributes of a speaking community negatively influence how a speaker is perceived, often based on completely imagined characteristics. In one example of this stereotyping, Rubin (1992) showed that native-speaking English listeners perceived an audio lecture paired with an image of a Chinese-looking female to be heavily accented, so much so that it interfered with the understanding of the lecture content. When the same audio was presented alongside an image of a Caucasian female, however, the content was understood significantly better and the speaker was rated as less accented, even though the audio was recorded, in both instances, by the same native English speaker from Ohio. Such preconceptions have also been found to affect the assessment of student work. For example, teachers of grade 3 and 4 schoolchildren provided higher evaluations of students' writing when it was paired with the speech of native English speakers than when students' writing was presented along with the speech of Spanish-accented speakers, regardless of which students actually produced the written work (Ford, 1984).

2.2 Manipulating Social Bias

Considering that listener-based evaluations of L2 speech are common in both research (e.g., Derwing & Munro, 2015) and assessment (e.g., Harding, 2012; Isaacs, 2013) contexts, it is important to thoroughly explore not only the effects of social biases on listener attitudes, but also to understand the conditions under which biases are amplified, as well as how biases can be mitigated toward better rating stability. For example, stereotypical yet extraneous information provided to participants can affect their performance. In a study conducted in an Italian-German bilingual community in Italy, Paladino et al. (2009) found that when Italian participants were simply reminded of the widely held perception that Italians of the region were known to have poor ability in German, they underperformed in their L2 German oral and written tests. Along the same lines, American listeners in Niedzielski's (1999) study perceived vowel sounds differently, even though they were produced by the same speaker, simply based on information that was provided prior to the listening task about the speaker being a resident of either the United States or Canada.

It is also possible to minimize existing biases. For instance, it is well documented that negative attitudes exist among native-speaking university students toward accented international teaching assistants (Halleck, 2008) and instructors (Hertel & Sunderman, 2009). Here, explicit training—in the form of increased exposure—can be useful in reducing bias. For instance, Staples et al. (2014) involved native-speaking undergraduate students in informal, cooperative contact activities with L2 speakers for eight weeks. Students who engaged in contact with L2 speakers subsequently rated L2 instructors more favorably on measures of accentedness (how closely the speaker approximates the target language variety), comprehensibility (how easy the speaker is to understand), and overall teaching ability than the group that had not

taken part in additional cooperative tasks. Beyond explicit instruction, other mitigation tactics have also been successful in influencing socially constructed biases toward L2 speech. For instance, Hansen et al. (2014) invited German-speaking raters to walk in the shoes of an L2 speaker by conversing with a confederate researcher in the participants' own L2 (English) prior to rating the L2 German speech of Turkish speakers. Those who engaged in this form of perspective taking assigned higher ratings to the L2 speakers than raters who did not activate their L2 prior to the session.

3 Exploring Bias

Given that listener-based evaluations of L2 speech are influenced by extraneous factors, such as social biases and stereotypical views, these evaluations may not be as stable as originally thought. Nevertheless, when it comes to the assessment of L2 speech, human ratings are essential (Derwing & Munro, 2015). Teachers regularly evaluate L2 speakers in low-stakes assessments, such as oral presentations and tests, and trained assessors evaluate L2 speakers in higher-stakes contexts, such as standardized examinations. Naïve (untrained) listeners are also called upon often to rate speech for various dimensions, including accentedness and comprehensibility (Derwing & Munro, 2009). It is further common for untrained listeners to provide evaluations of L2 speakers that extend beyond speech itself. For example, individuals with no training in speech assessment have been asked to judge L2 speakers' socioeconomic status (Deprez-Sims & Morris, 2010), educational achievement (Campbell-Kibler, 2007), and competence (Baquiran & Nicoladis, 2020). Such evaluations often have implications for future work and study opportunities, wages, and the quality of healthcare that a person receives (Halim et al., 2017; Timming, 2017). If speech ratings are susceptible to social influences, then it is crucial to seek a better understanding of ways in which such rating-irrelevant variance (i.e., extraneous, uncontrolled variables that can influence assessment) can be minimized.

3.1 Social Bias and Naïve Listeners' Evaluations of L2 Speech

The initial objective of our work was to determine the effect of deliberate positive and negative social bias manipulation on naïve (untrained) listeners' ratings of L2 speech (Taylor Reid et al., 2019). This study, which explored the effects of social bias on listeners' assessments of native French speakers of L2 English from Quebec, Canada, was fueled by the idea that social influences, however slight, prior to the rating session might sway ratings assigned by naïve listeners relative to listeners who had

not experienced any social influence prior to rating, especially in a social environment with a history of tension between English- and French-speaking communities.

Because the impact of social bias on listeners' judgments of L2 speech is likely determined by their specific experiences (e.g., Kang & Rubin, 2009; Wigboldus et al., 2005), the listeners recruited for this study represented a broad age range. We expected to find that older and younger listeners might differ in their ratings according to the impact of Quebec's language policy on their respective generations. The 1977 French Language Charter (Bill 101), which designated French as the sole official language of Quebec and restricted the use of English in public domains (including education) as a way of strengthening the ethnolinguistic vitality of francophones in Quebec (Corbeil, 2007), was expected to have heavily influenced the attitudes of older (40+) listeners. These listeners would have been children or young adults when the status of English changed from majority to minority, which would make them particularly sensitive to issues affecting English-speaking Quebecers. In contrast, younger listeners would have been raised and schooled at a time when the official status of French had been less contested, making them less sensitive to English-centered social influences. We therefore predicted that any social bias in listener evaluations would be qualified by listeners' age, leading to a more pronounced bias among older rather than younger listeners.

Sixty listeners were randomly assigned to one of three groups that engaged in a similar rating task, except for a brief personal story that a researcher shared with the listeners in two of the three groups at the outset of the session. In the negative manipulation group, 20 listeners (ages 19–66) heard negative comments by the researcher about a recent encounter with an L2 English speaker. In the positive manipulation group, 20 listeners (ages 18–72) heard a comparable opinion of the same length and emotional content reflecting the researcher's positive experience with an L2 English speaker. The 20 listeners (ages 20–65) in the baseline group rated the speech samples without any such manipulation. Regardless of group assignment, all listeners heard the same 40 brief L2 English narratives recorded by native French speakers from Quebec, assessing each speaker for accentedness, comprehensibility, and flow (overall pacing and speed of utterance delivery), as well as for specific pronunciation issues, such as segmental errors (accuracy in articulation of consonants and vowels) and intonation (natural rise and fall in pitch).

When the listeners were exposed to a positive bias manipulation, they generally behaved similarly in their ratings, irrespective of their age. Younger listeners upgraded the speakers for four of the five targeted measures (accentedness, comprehensibility, intonation, flow), while older listeners enhanced the speakers' evaluations for two measures (comprehensibility, intonation), compared to baseline listeners' assessments. However, the rating behaviors of the younger and older listeners diverged under a negative bias manipulation. Negatively oriented younger listeners pushed back against the researcher's negative comments, providing more favorable ratings for all five measures. However, this was not the case for the negatively oriented older listeners, who downgraded the same speakers relative to baseline listeners' evaluations. To put it another way, the researcher's biased comments about an L2 English

speaker upset rating stability for the older listeners, who went along with the positivity (upgrading the speakers) and the negativity (downgrading the speakers), but the same comments functioned as an unexpected positive stimulus for the younger listeners under both the negative and positive bias manipulation conditions.

3.2   Social Bias and Teachers' Evaluations of L2 Speech

Armed with a clearer understanding of the influence of social bias on naïve listeners' evaluations of L2 speech, we then sought to investigate how that influence would affect expert raters in a new linguistic environment, this time with language teachers evaluating their students' performance (Taylor Reid et al., 2020). We specifically examined whether teachers of L2 German—as evaluators of their students' speaking performance—might also be sensitive to a social bias manipulation, just like the naïve listeners in our earlier work. We additionally investigated whether sensitivity to social bias might differ for teachers who are themselves either native speakers or non-native speakers of the language they are teaching.

With respect to native versus non-native teacher differences, it was possible that both native and non-native teachers of German would provide similar ratings for L2 speakers of German, regardless of the biasing orientation (e.g., Crowther et al., 2016; Derwing & Munro, 2013), because they are all part of the same professional group. However, negative comments made about L2 speakers might particularly resonate with the non-native teachers, invoking feelings of empathy that could carry over to more generous ratings (e.g., Hansen et al., 2014). Alternatively, the non-native teachers might show more negativity in their evaluations when exposed to social bias, given that L2 listeners sometimes provide harsher evaluations than native listeners (e.g., Kang, 2012; Rose, 2017; Rossiter, 2009).

In this study, conducted in the context of teaching and learning German as a foreign language in the English-speaking province of Alberta, Canada, we asked teachers of German to evaluate the speech of 24 intermediate to advanced L2 German speakers for the same five measures (i.e., accentedness, comprehensibility, flow, segmental errors, and intonation) as in our earlier study. Because teachers of L2 German in Alberta represent a relatively small group, which made it difficult to carry out a large-scale study, we recruited two comparable groups of teachers: (a) an experimental group made up of seven non-native and seven native teachers of German, and (b) a control group made up of seven non-native and seven native teachers of German. As in our earlier work, the two groups of teachers engaged in the same rating task, but only the experimental group received negative bias manipulation—that is, the researcher provided a negative opinion about the L2 German skills of a hypothetical learner of German prior to asking the teachers to rate the L2 German speech samples.

The native and non-native teachers provided comparable ratings of the intonation and fluency of L2 German speech, demonstrating similar rating behaviors for these

speech dimensions. However, in response to negative bias, the native and non-native teachers diverged in their evaluations of the remaining three dimensions (accentedness, comprehensibility, and segmental errors). The native teachers downgraded the performance of L2 speakers, which corresponded to medium-strength statistical effect. In contrast, the non-native teachers provided more favorable evaluations for the same speakers. Put differently, the native teachers who heard negative comments about L2 students' German appeared to go along with the negativity, downgrading the speakers, whereas the non-native teachers seemed to show (enhanced) empathy with fellow L2 speakers, upgrading their ratings.

3.3 Summary and Outlook

Taken together, these two investigations shed new light on the stability of L2 speech ratings among both naïve and expert raters. In Taylor Reid et al. (2019), we found strong, consistent effects of positive and negative bias manipulations on all five targeted speech measures, such that the ratings provided by listeners under social bias diverged significantly from the ratings provided by baseline listeners. These findings add to the growing body of research in applied linguistics (e.g., Winke et al., 2013) and social psychology (e.g., Paladino et al., 2009) targeting various sources of bias in measures of L2 learning and use, and invite further investigations into social, attitudinal, and emotional underpinnings of listener assessments of L2 speech.

Our findings also cast doubt on the relative stability of human ratings of L2 speech, pointing to the importance of social context, defined both narrowly (as an immediate rating situation) and broadly (as a sociopolitical environment). Although the local context was tightly controlled, in that it took place in a laboratory, the real-world settings in which assessors find themselves before they pass judgment on L2 speakers are less rigidly controlled, which makes speech ratings particularly susceptible to various social influences. In fact, people are often unaware of the experiences that activate their preconceived ideas or stereotypes (Molden, 2014), so any negativity overtly or covertly attributed to L2 speakers can result in behaviors with important real-life consequences that extend, for example, to employment, wages, and healthcare (e.g., Halim et al., 2017; Hansen & Dovidio, 2016; Timming, 2017).

These concerns certainly extend beyond contexts involving naïve listeners. In Taylor Reid et al. (2020), we showed that language teachers are influenced by social bias, which differed in nontrivial ways in its influence on native versus non-native teachers. Language teachers evaluate L2 learners on a regular basis. Sometimes they carry out high-stakes evaluations that may determine, for example, whether a learner may study or work in a given target language setting. In such instances, it is essential that teachers be aware of both their own biases (as native speakers of the target language or as fellow non-native speakers) and the ways in which their assessments may be affected by comments provided by others. Importantly, too, employing multiple raters to carry out assessments in high-stakes settings may safeguard against the biases of individual raters. A prudent take-home message arising

from this research is that not only are ratings more malleable than previously thought, but neither naïve listeners nor language teachers, as members of their respective sociolinguistic groups—whether they are laypersons or experienced language teachers, whether they are young or old, or whether they are native or non-native speakers themselves—are immune to social biases.

4 Mitigating Social Bias

Having established (at least some) effects of social bias on L2 speech ratings, we then explored ways to minimize rating-irrelevant social influences on evaluations of L2 speech. One mitigation strategy involves awareness raising through perspective taking (Boland & Tenkasi, 1995). Perspective taking refers to various activities whose goal is to guide people to consider various facets of another individual with whom they might surprisingly share commonalities. For instance, in a study of English-speaking university students' evaluations of L2 speakers, before eliciting speech judgments, Weyant (2007) asked some students to write about a day in the life of an L2 speaker, while other students were given no such instructions. The students who took the L2 speaker' perspective assigned her higher ratings of ability and accomplishment, compared to those who were not asked to write from the speaker's perspective. In another example, Zhang (2017) immersed pre-service music teachers in a 20-minute music class with all instructions and content delivered in Mandarin as a way of approximating the experience of learning the course content in another language. The teachers subsequently reported feeling anxious, confused, and frustrated, revealing an emotional response to perspective taking that may have allowed them to develop a new understanding of L2 speakers and their challenges.

4.1 Task Practice as Mitigation Strategy

Given that various forms of perspective taking appear to be successful at reducing listener bias (Weyant, 2007) or in creating greater awareness of the individuals being evaluated (Zhang, 2017), we reasoned that engaging listeners in task practice—essentially by asking them to perform the same speaking task as the speakers to be assessed—could stabilize listener behaviors by reducing social bias effects on their ratings. We explored this hypothesis in Taylor Reid et al. (2021).

As a starting point, we considered our finding from Taylor Reid et al. (2019): young English-French bilingual listeners (all dominant in English) were susceptible to negative and positive social bias in evaluation of L2 English speech. As discussed previously, the younger listeners exposed to both negative and positive bias manipulations upgraded L2 speakers significantly in their ratings, compared to the evaluations by those who were not exposed to a biasing social commentary, demonstrating rating-irrelevant variance in their speech assessments. With this finding in mind, in

our most recent study (Taylor Reid et al., 2021), we targeted similar English-French bilinguals to examine whether engaging them in task practice in their more versus less dominant language could reduce rating-irrelevant variance in speech ratings arising through social bias. Because all social bias effects among younger listeners in our previous study led to inflated (more positive) ratings, we anticipated that task practice in listeners' more dominant language (English) would be more effective at minimizing social bias effects (i.e., bringing listeners' ratings more in line with the baseline listeners' ratings), compared to task practice in their less dominant language (French). In other words, those who are called upon to use their more dominant language prior to engaging in speech ratings might have higher expectations of L2 speakers, resulting in a reduction of any leniency that might be brought about by a biasing commentary. In contrast, the use of a less dominant L2 might be associated with rating leniency (e.g., Weyant, 2007), which may not be as effective at minimizing rating-irrelevant variance that has already been amplified through social bias.

Similar to the listeners in our earlier work, the 70 young bilingual listeners in this study were exposed to positively or negatively worded biasing comments. Most critically, however, before they provided their assessments of L2 speakers' accentedness and comprehensibility, 20 participants completed the same narrative task as the L2 speakers in their stronger language (English); another 20 participants performed the same task in their weaker language (French). The remaining 30 participants in the baseline groups did not engage in task practice and completed ratings of the 40 speech samples with negative, positive, or no social bias imposed. Only English task practice appeared to significantly reduce rating-irrelevant effects of social bias on listener assessments, and only under negative bias manipulation. In other words, those listeners who practiced the task in English were less likely to upgrade the speech of L2 speakers under the negative bias manipulation.

We reasoned that engaging the listener in the same speaking task completed by L2 speakers may have encouraged perspective taking in similar ways to writing about the life of a L2 speaker (Weyant, 2007). Performing the task in English rather than French was also useful, as it provided the listeners with a model to use in their evaluation of the speakers' L2 English speech, which is consistent with the positive role of increased task familiarity in rater training (Davis, 2016). In essence, increased task familiarity, along with realistic performance expectations available to the listeners through English task practice, may have limited the impact of the (negative) biasing commentary. A preliminary take-home message here is that task practice might be a medium through which L2 accentedness and comprehensibility ratings can be stabilized, thereby countering the effects of social bias.

4.2 Next Steps

In light of these findings, future work investigating human ratings of L2 speech should consider the impact of additional task practice and perspective taking interventions on listener-assessed dimensions of L2 speech. These could include reading or listening to anecdotes about situations where L2 speakers experienced prejudice or enhanced empathy from their interlocutors on the basis of their speech. Similarly, listeners could be asked to comment on situations in their own lives in which they experienced prejudice or enhanced empathy on the basis of their language performance or factors unrelated to their linguistic competence. If such interventions are effective in terms of encouraging more positive assessments of L2 speech in a laboratory setting, they could be utilized in the training of individuals tasked with assessing L2 speakers on a regular basis. For example, L2 teachers and examiners could be encouraged to complete tasks similar to those they are assessing before they begin their assessments.

Along similar lines, one aspect of training provided to human resources personnel could involve roleplaying that teases apart linguistic issues associated with people's speech from other factors related to their professional competence. In even higher-stakes contexts where assessment of credibility is paramount—such as those involving legal interaction with L2 speakers (as part of traffic stops, border crossings, courtroom proceedings)—efforts could move beyond conventional tactics aimed at increased understanding (e.g., diversity training) to those that might expand the mindset of the participant by combining perspective taking with other successful interventions such as intercultural communication opportunities, as can be achieved, for example, through virtual reality (see Salmanowitz, 2016).

5 Broader Implications, Future Work, and Conclusions

Collectively, our recent findings have enabled us to gain a clearer picture of how some forms of social bias, such as positively or negatively worded comments, and some types of interventions, such as task practice, can affect rater behavior, but tough work remains if we are to apply this knowledge to real-world contexts. For instance, mitigation found to be effective in a laboratory setting might naturally be extended to pedagogical contexts, where greater consistency and fairness in evaluations of L2 speech might be ensured through rater training, task familiarity exercises, and awareness-raising tactics that call attention to subconsciously held biases. The same findings might be applied to other organized contexts, such as the courtroom or corporate environment, where one might lean on diversity training and other tactics that guide decision-makers toward effective disentanglement of an L2 speakers' character and credibility from their linguistic status.

But what of the less-organized day-to-day interactions: traveler to traveler, teammate to teammate, neighbor to neighbor? To effect change in such circles, a more

thorough exploration of the role of contextual factors—such as linguistic environment and exposure—on social attitudes toward L2 speech is warranted. We might also seek a sharpened understanding of how to best explain negative attitudes from a theoretical perspective so that we can maximize our efforts toward neutralizing negative attitudes at their roots. This could involve further examination through the lenses of such frameworks as communication accommodation and intergroup contact (Berry, 1997; Bourhis et al., 2012; Dragojevic et al., 2016). For instance, even if people enter a conversation with the best of intentions to collaborate toward effective and equitable communication, being able to control deep-seated biases and resist the urge to fall back on the upkeep and affirmation of one's social identity—often to the detriment of the interlocutor—is another matter.

More importantly, long-term mitigation of biases is likely dependent on targeted interaction. When the speaker and listener are separated—as in laboratory rater studies, scoring of university placement tests, or even listening to voicemail messages from prospective employees—there is no opportunity for interaction to occur. Furthermore, increased exposure to ethnic, linguistic, and cultural diversity might alleviate at least some negative biases in areas where regular contact with L2 speakers is limited. Such a finding could provide an indication that an individual's social network might be a contributing factor. If so, an investigation of community engagement opportunities that bring culturally diverse groups together in an appealing social context might be an excellent form of real-world mitigation. Additional research and practice might then focus on how effective such efforts are at permanently altering attitudes toward more equitable treatment of L2 speakers across contexts.

References

Babel, M., & Russell, J. (2015). Expectations and speech intelligibility. *The Journal of the Acoustical Society of America, 137*(5), 2823–2833. https://doi.org/10.1121/1.4919317

Baquiran, C. L. 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

Berry, J. W. (1997). Immigration, acculturation, and adaptation. *Applied Psychology: An International Review, 46*(1), 5–34. https://doi.org/10.1111/j.1464-0597.1997.tb01087.x

Boland, R. J., & Tenkasi, R. V. (1995). Perspective making and perspective taking in communities of knowing. *Organization Science, 6*(4), 350–372. https://doi.org/10.1287/orsc.6.4.350

Bourhis, R. Y., Sioufi, R., & Sachdev, I. (2012). Ethnolinguistic interaction and multilingual communication. In H. Giles (Ed.), *The handbook of intergroup communication* (pp. 100–115). Routledge. https://doi.org/10.4324/9780203148624.ch8

Brennan, E. M., & Brennan, J. S. (1981). Accent scaling and language attitudes: Reactions to Mexican-American English speech. *Language and Speech, 24*(3), 207–221. https://doi.org/10.1177/002383098102400301

Campbell-Kibler, K. (2007). Accent, (ING), and the social logic of listener perceptions. *American Speech, 82*(1), 32–64. https://doi.org/10.125/00031283-2007-002

Corbeil, J. C. (2007). *L'embarras des langues: Origine, conception et évolution de la politique linguistique québécoise*. Québec Amérique.

Crowther, D., Trofimovich, P., & Isaacs, T. (2016). Linguistic dimensions of second language accent and comprehensibility. *Journal of Second Language Pronunciation, 2*(2), 160–182. https://doi.org/10.1075/jslp.2.2.02cro

Dalton-Puffer, C., Kaltenboeck, G., & Smit, U. (1997). Learner attitudes and L2 pronunciation in Austria. *World Englishes, 16*(1), 115–128. https://doi.org/10.1111/1467-971X.00052

Davis, L. (2016). The influence of training and experience on rater performance in scoring spoken language. *Language Testing, 33*(1), 117–135. https://doi.org/10.1177/0265532215582282

Deprez-Sims, A. S., & Morris, S. B. (2010). Accents in the workplace: Their effects during a job interview. *International Journal of Psychology, 45*(6), 417–426. https://doi.org/10.1080/00207594.2010.499950

Derwing, T. M., & Munro, M. J. (2009). Comprehensibility as a factor in listener interaction preferences: Implications for the workplace. *Canadian Modern Language Review, 66*(2), 181–202. https://doi.org/10.3138/cmlr.66.2.181

Derwing, T. M., & Munro, M. J. (2013). The development of L2 oral language skills in two L1 groups: A 7-year study. *Language Learning, 63*(2), 163–185. https://doi.org/10.1111/lang.12000

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research.* John Benjamins. https://doi.org/10.1075/lllt.42

Dragojevic, M., Gasiorek, J., & Giles, H. (2016). Communication accommodation theory. In C. R. Berger & M. E. Roloff (Eds.), *The international encyclopedia of interpersonal communication* (pp. 1–20). Wiley. https://doi.org/10.1002/9781118540190.wbeic006

Ford, C. E. (1984). The influence of speech variety on teachers' evaluation of students with comparable academic ability. *TESOL Quarterly, 18*(1), 25–40. https://doi.org/10.2307/3586333

Giles, H., & Watson, B. (Eds.). (2013). *The social meanings of language, dialect, and accent: International perspectives on speech styles.* Peter Lang. https://www.peterlang.com/document/1109208

Gluszek, A., & Dovidio, J. F. (2010). The way they speak: A social psychological perspective on the stigma of non-native accents in communication. *Personality and Social Psychology Review, 14*(2), 214–237. https://doi.org/10.1177/1088868309359288

Halim, M. L., Moy, K. H., & Yoshikawa, H. (2017). Perceived ethnic and language-based discrimination and Latina immigrant women's health. *Journal of Health Psychology, 22*(1), 68–78. https://doi.org/10.1177/1359105315595121

Halleck, G. B. (2008). The ITA problem: A ready-to-use simulation. *Simulation & Gaming, 39*(1), 137–146. https://doi.org/10.1177/1046878107308060

Hansen, K., & Dovidio, J. F. (2016). Social dominance orientation, nonnative accents, and hiring recommendations. *Cultural Diversity and Ethnic Minority Psychology, 22*(4), 1–8. https://doi.org/10.1037/cdp0000101

Hansen, K., Rakic, T., & Steffens, M. C. (2014). When actions speak louder than words: Preventing discrimination of nonstandard speakers. *Journal of Language and Social Psychology, 33*(1), 68–77. https://doi.org/10.1177/0261927X13499761

Harding, L. (2012). Pronunciation assessment. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1–6). Wiley-Blackwell. https://doi.org/10.1002/9781405198431.wbeal0966

Heaton, H., & Nygaard, L. C. (2011). Charm or harm: Effect of passage content on listener attitudes toward American English accents. *Journal of Language and Social Psychology, 30*(2), 202–211. https://doi.org/10.1177/0261927X10397288

Hertel, T. J., & Sunderman, G. (2009). Student attitudes toward native and non-native language instructors. *Foreign Language Annals, 42*(3), 468–482. https://doi.org/10.1111/j.1944-9720.2009.01031.x

Isaacs, T. (2013). Assessing pronunciation. In A. J. Kunnan (Ed.), *The companion to language assessment* (pp. 140–155). Wiley-Blackwell. https://doi.org/10.1002/9781118411360.wbcla012

Kagedan, A. L. (2020). *The politics of othering in the United States and Canada.* Palgrave Macmillan. https://doi.org/10.1007/978-3-030-52444-9

Kang, O. (2012). Impact of rater characteristics and prosodic features of speaker accentedness on ratings of international teaching assistants' oral performance. *Language Assessment Quarterly, 9*(3), 249–269. https://doi.org/10.1080/15434303.2011.642631

Kang, O., & Rubin, D. L. (2009). Reverse linguistic stereotyping: Measuring the effect of listener expectations on speech evaluation. *Journal of Language and Social Psychology, 28*(4), 441–456. https://doi.org/10.1177/0261927X09341950

Krahn, H., Derwing, T., Mulder, M., & Wilkinson, L. (2000). Educated and underemployed: Refugee integration into the Canadian labour market. *Journal of International Migration and Integration/Revue de l'integration et de la migration internationale, 1*(1), 59–84. https://doi.org/10.1007/s12134-000-1008-2

Labov, W. (1972). On the mechanism of language change. In J. J. Gumperz & D. Hymes (Eds.), *Directions in sociolinguistics* (pp. 312–338). Holt, Rinehart, and Winston.

Lalwani, A. K., Lwin, M., & Li, K. L. (2005). Consumer responses to English accent variations in advertising. *Journal of Global Marketing, 18*(3–4), 143–165. https://doi.org/10.1300/J042v18n03_07

Lindemann, S. (2002). Listening with an attitude: A model of native speaker comprehension of nonnative speakers in the United States. *Language in Society, 31*(3), 419–441. https://doi.org/10.1017/S0047404502020286

Lindemann, S., & Subtirelu, N. (2013). Reliably biased: The role of listener expectation in the perception of second language speech. *Language Learning, 63*(3), 567–594. https://doi.org/10.1111/lang.12014

Matsuda, M. J. (1991). Voices of America: Accent, antidiscrimination law, and a jurisprudence for the last reconstruction. *Yale Law Journal, 100*(5), 1329–1407. https://doi.org/10.2307/796694

Molden, D. C. (2014). Understanding priming effects in social psychology: What is "social priming" and how does it occur? *Social Cognition, 32,* 1–11. https://doi.org/10.1521/soco.2014.32.supp.1

Mullin, B. A., & Hogg, M. A. (1999). Motivations for group membership: The role of subjective importance and uncertainty reduction. *Basic and Applied Social Psychology, 21*(2), 91–102. https://doi.org/10.1207/S15324834BA210202

Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology, 18*(1), 62–85. https://doi.org/10.1177/0261927X99018001005

Paladino, M.-P., Poddesu, L., Rauzi, M., Vaes, J., Cadinu, M., & Forer, D. (2009). Second language competence in the Italian-speaking population of Alto Adige/Südtirol: Evidence for linguistic stereotype threat. *Journal of Language and Social Psychology, 28*(3), 222–243. https://doi.org/10.1177/0261927X09335333

Rose, R. L. (2017). Differences in second language speech fluency ratings: Native versus nonnative listeners. In *Proceedings of the international conference "Fluency & Disfluency Across Languages and Language Varieties"* (pp. 101–103). Catholic University of Louvain. http://hdl.handle.net/2078.1/195807

Rossiter, M. J. (2009). Perceptions of L2 fluency by native and non-native speakers of English. *Canadian Modern Language Review, 65*(3), 395–412. https://doi.org/10.3138/cmlr.65.3.395

Rubin, D. L. (1992). Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. *Research in Higher Education, 33*(4), 511–531. https://doi.org/10.1007/BF00973770

Ryan, E. B. (1983). Social psychological mechanisms underlying native speaker evaluations of non-native speech. *Studies in Second Language Acquisition, 5*(2), 148–159. https://doi.org/10.1017/S0272263100004824

Ryan, E. B., Carranza, M. A., & Moffie, R. W. (1977). Reactions toward varying degrees of accentedness in the speech of Spanish-English bilinguals. *Language and Speech, 20*(3), 267–273. https://doi.org/10.1177/002383097702000308

Salmanowitz, N. (2016). Unconventional methods for a traditional setting: The use of virtual reality to reduce implicit racial bias in the courtroom. *The University of New Hampshire Law Review, 15*(1), 118–160. https://scholars.unh.edu/unh_lr/vol15/iss1/2/

Staples, S., Kang, O., & Wittner, E. (2014). Considering interlocutors in university discourse communities: Impacting U.S. undergraduates' perceptions of ITAs through a structured contact program. *English for Specific Purposes*, *35*, 54–65. https://doi.org/10.1016/j.esp.2014.02.002

Stroebe, W., & Insko, C. A. (2013). Stereotype, prejudice, and discrimination: Changing conceptions in theory and research. In D. Bar-Tal, C. F. Graumann, A. W. Kruglanski, & W. Stroebe (Eds.), *Stereotyping and prejudice: Changing conceptions* (pp. 3–36). Springer. https://doi.org/10.1007/978-1-4612-3582-8_1

Taylor Reid, K., O'Brien, M. G., Trofimovich, P., & Bajt, A. (2020). Testing the malleability of teachers' judgments of second language speech. *Journal of Second Language Pronunciation, 6*(2), 237–265. https://doi.org/10.1075/jslp.19015.tay

Taylor Reid, K., Trofimovich, P., & O'Brien, M. G. (2019). Social attitudes and speech ratings: Effects of positive and negative bias on multiage listeners' judgments of second language speech. *Studies in Second Language Acquisition, 41*(2), 419–442. https://doi.org/10.1017/S0272263118000244

Taylor Reid, K., Trofimovich, P., & O'Brien, M. G., & Tsunemoto, A. (2021). Using task practice to reduce social influences on listener evaluations of second language accent and comprehensibility. *International Journal of Listening*. Advance online publication. https://doi.org/10.1080/10904018.2021.1904933

Timming, A. R. (2017). The effect of foreign accent on employability: A study of the aural dimensions of aesthetic labour in customer-facing and non-customer-facing jobs. *Work, Employment and Society, 31*(3), 409–428. https://doi.org/10.1177/0950017016630260

Weyant, J. M. (2007). Perspective taking as a means of reducing negative stereotyping of individuals who speak English as a second language. *Journal of Applied Social Psychology, 37*(4), 703–716. https://doi.org/10.1111/j.1559-1816.2007.00181.x

Wigboldus, D. H., Spears, R., & Semin, G. R. (2005). When do we communicate stereotypes? Influence of the social context on the linguistic expectancy bias. *Group Processes & Intergroup Relations, 8*(3), 215–230. https://doi.org/10.1177/1368430205053939

Winke, P., Gass, S., & Myford, C. (2013). Raters' L2 background as a potential source of bias in rating oral performance. *Language Testing, 30*(2), 231–252. https://doi.org/10.1177/0265532212456968

Yzerbyt, V. Y., Schadron, G., Leyens, J. P., & Rocher, S. (1994). Social judgeability: The impact of metainformational cues on the use of stereotypes. *Journal of Personality and Social Psychology, 66*(1), 48–55. https://doi.org/10.1037/0022-3514.66.1.48

Zhang, Y. (2017). Walking a mile in their shoes: Developing pre-service music teachers' empathy for ELL students. *International Journal of Music Education, 35*(3), 425–434. https://doi.org/10.1177/0255761416647191

Zschirnt, E., & Ruedin, D. (2016). Ethnic discrimination in hiring decisions: A meta-analysis of correspondence tests 1990–2015. *Journal of Ethnic and Migration Studies, 42*(7), 1115–1134. https://doi.org/10.1080/1369183X.2015.1133279

Kym Taylor Reid is an Applied Linguistics Ph.D. candidate in the Department of Education at Concordia University and a part-time faculty member at McGill University in Montreal, Canada. Her research focuses on social biases against non-native speech, social network analysis, sociolinguistic aspects of second language acquisition, and the teaching of second language pronunciation.

Pavel Trofimovich is Professor of Applied Linguistics in the Department of Education at Concordia University, Montreal, Canada. His research focuses on cognitive aspects of second language processing, second language phonology, sociolinguistic aspects of second language acquisition, and the teaching of second language pronunciation.

Mary Grantham O'Brien is Professor of German in the School of Languages, Linguistics, Literatures and Cultures at the University of Calgary, Calgary, Canada. Her research focuses on the perception and production of second language speech with an emphasis on suprasegmentals.