Using task practice to reduce social influences on listener evaluations of second language accent and comprehensibility

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

Abstract

This study examined whether task practice can reduce social influences (operationalized in terms of positive or negative social priming) on listener-based evaluations of second language (L2) speech. Seventy English–French bilingual listeners evaluated audio-recorded picture narratives made by 40 French speakers of L2 English for accent and comprehensibility. Before rating the narratives, 40 listeners engaged in task practice by completing the same speech production task as the L2 speakers in English (20) or French (20) and then experienced a social priming intervention (10 negative, 10 positive per group). An additional 30 baseline listeners evaluated the same narratives without engaging in any task practice and with either negative (10), positive (10), or no social priming (10) imposed. English task practice mitigated social influences on listener assessments of both accent and comprehensibility, but only under negative priming. Engaging listeners in task practice, particularly in their stronger language, can reduce social influences on listener-based assessments of L2 speech.

Keywords: task practice, social influence, accent, comprehensibility, second language

Using task practice to reduce social influences on listener evaluations of second language accent and comprehensibility

Listeners are regularly called upon to make judgments about second language (L2) speakers in high-stakes contexts, including employment (Timming, 2017), healthcare (Halim et al., 2017), and education (Winke et al., 2013). Prior research has enlisted both expert raters (e.g., teachers, linguists) and naïve listeners to evaluate L2 speakers for multiple rated dimensions, including accent, which captures how closely the speaker approximates the target language variety, and comprehensibility, which encompasses how easily listeners understand the speaker (Derwing & Munro, 2015). Often, untrained listeners are also tasked with evaluating L2 speakers along dimensions that extend beyond speech, including perceived socioeconomic status (Deprez-Sims & Morris, 2010), educational achievement (Campbell-Kibler, 2007), and competence (Baquiran & Nicoladis, 2020). Given that speech-based judgments have high-stakes implications for L2 speakers (Halim et al., 2017; Timming, 2017), it is critical to explore how listeners' evaluations of L2 speakers might be affected by various social or contextual influences which are irrelevant to L2 speakers' performance but might nevertheless impact listener assessments. The goal of this study was, therefore, to document some of these social influences, exploring the effect of exposing listeners to positive or negative opinions prior to the rating task, and to examine ways to minimize the impact of these influences on listeners' evaluations of L2 speech.

Although previous research has often reported high levels of listener agreement for evaluations of L2 speech along many continua (Plonsky & Derrick, 2016), the validity of speech ratings has been called into question (e.g., Kang & Rubin, 2009; Taylor Reid et al., 2019). At the root of this issue is the observation that human decisions are affected, both positively and negatively, by multiple outside social variables (Haladyna & Downing, 2004). These variables

have been discussed in prior research using such terms as biases, stereotype threats, and attitudes, but here we specifically operationalize them in terms of researcher- or task-imposed positive or negative social influences on listeners. The vast majority of socially-engendered influences on listeners can be understood within the broad construct of social priming, which refers to people's exposure to information that can potentially activate their social views (e.g., beliefs, stereotypes) and thus influence people's reactions to situations or stimuli (Molden, 2014). One noteworthy aspect of social priming is that people might not only succumb to positive or negative social influences, but that they might also act in accordance with the beliefs or stereotypes activated through priming (Dijksterhuis & Bargh, 2001). Various effects compatible with social priming include contextual influences, such as when juror decisions are impacted by extralegal information, including pretrial publicity (Kerr et al., 1999). Similar effects involve identity priming, where biracial individuals appear to act in accordance with either Black or White identity depending on whether they are led to think about their Black or White heritage (Gaither et al., 2015), and stereotype threat, where exposure to a negative stereotype about an ethnolinguistic group leads to speaker underperformance in language proficiency tests (Paladino et al., 2009). In our own prior work, we found that a social prime (delivered through an anecdote about L2 speakers' language skills) resulted in inflated ratings of L2 speech among naïve listeners, compared to the ratings provided by those not exposed to a priming manipulation (Taylor Reid et al., 2019).

L2 speech researchers have introduced a range of intervention strategies to counter at least some influences on listener-based evaluations of L2 speech, including those arising from social priming. One strategy involves instruction or exposure interventions. To reduce stereotype-driven evaluations of L2 speakers, Staples et al. (2014) investigated the effectiveness of

informal, cooperative contact activities involving native-speaking university students and L2 speakers. After eight weeks of contact activities, students in the contact group rated L2 instructors on their university campus as being less accented, more comprehensible, and as having greater teaching ability, compared to students who did not participate in any contact activities. In another intervention study involving listener training, Derwing et al. (2002) familiarized students in a social work class with Vietnamese-accented L2 English speech, engaged them in cross-cultural discussion, and provided instruction on specific pronunciation targets. Although the intervention did not improve students' understanding of Vietnamese-accented speech, it did improve students' self-reported confidence in rating.

Another strategy aimed at reducing social influences on listener-based assessments of L2 speech involves awareness raising through perspective taking. Perspective taking—a way to guide people to consider various facets of another individual with whom they might surprisingly share commonalities—can be particularly useful when trying to tackle socially-engendered behaviors (Boland & Tenkasi, 1995). In a study of undergraduate students' perceptions of minorities (e.g., the elderly, African Americans), Galinsky and Moskowitz (2000) showed that perspective taking through writing about a day in the life of a minority individual decreased negative stereotyping and increased awareness of similarities between the listener and the individual. In another study, Weyant (2007) asked 160 native English-speaking college students to listen to a recording of either a native speaker or an L2 speaker. Some participants were then asked to write about a day in the life of that speaker, while the others were given no such instruction. The students who took the perspective of the L2 speaker assigned her higher ratings of ability and accomplishment, compared to those who were not asked to write from the speaker's perspective. These tactics need not involve lengthy interventions or activities. Hansen

et al. (2014) engaged half of their German listeners in a brief conversation in English (their L2) before they rated ethnic Turkish speakers' German speech. The listeners who essentially "stepped into the shoes" of L2 speakers by holding a brief conversation in their L2 tended to assign higher comprehensibility and employability ratings to Turkish-accented German speakers, compared to those who did not use their L2 before rating.

The Current Study

If listener-based assessments of L2 speech are susceptible to social influences, including those compatible with social priming effects, then it is crucial to seek a better understanding of ways in which these influences can be minimized, particularly for such key listener-evaluated dimensions as L2 accent and comprehensibility. Often targeted in high- and low-stakes assessments (Derwing & Munro, 2015), ratings of accent and comprehensibility capture the essential aspects of L2 speakers' performance relevant to L2 speakers and their interlocutors. Thus, this study's objective was to extend our prior work on social priming (Taylor Reid et al., 2019) by evaluating two types of task practice interventions designed to encourage perspective taking. The goal was to reduce social effects on English–French bilingual listeners evaluating French-accented L2 English speech. For one type of task practice, listeners engaged in a speech production task (i.e., the same task as the L2 speakers to be assessed) in their more dominant language (English); for the other type of practice, listeners performed this task in their less dominant language (French). The two types of task practice differed in the extent to which they encouraged listeners to adopt the perspective of L2 speakers. The English intervention allowed listeners to experience the same task completed by those rated. In contrast, the French intervention encouraged listeners not only to practice the same task, but also to experience being L2 speakers themselves.

As a starting point, we used our recent finding that young English–French bilingual listeners in Montreal (all dominant in English) are susceptible to negative and positive social priming in evaluation of French-accented L2 English speech (Taylor Reid et al., 2019). In that study, before providing speech ratings, listeners heard the researcher share a negative or positive anecdote (which delivered a negatively- or positively-charged prime). Listeners who were exposed to both the negative and the positive anecdote upgraded L2 speakers significantly in their ratings, compared to the evaluations by listeners who were not exposed to a priming anecdote. In essence, young English-dominant bilingual Montrealers (aged 18–40) went along with the positive but also "fought against" the negative, upgrading L2 speakers in their assessments in both cases (compared to baseline listeners' evaluations), thus revealing socially-based influences on their ratings.

To determine whether engaging listeners in task practice in their more versus less dominant language can reduce social influences on speech ratings arising through social priming, we recruited 40 English–French bilinguals, all representative of the same community of speakers as in the earlier study. Our assumption, based on our prior research, was that the effectiveness of task practice would depend on whether listeners share only a speaking task with L2 speakers (English task practice) or whether they share both a speaking task and a language status with L2 speakers by using a non-dominant, additional language (French task practice). Although there has been limited systematic investigation into social influences on L2 speech ratings, on the basis of the results of Hansen et al. (2014) and Weyant (2007), we expected task practice to minimize social priming effects on accent and comprehensibility, in the sense that task practice would bring speech ratings in line with the assessments by listeners not exposed to social priming. In addition, because all social priming effects in our prior work led to inflated (more positive)

ratings, we anticipated that task practice in listeners' more dominant language (English) would be more effective at minimizing social priming effects, thereby bringing all ratings in alignment with the baseline group's ratings, compared to task practice in listeners' less dominant language (French). Put differently, listeners 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 social priming. In contrast, the use of a less dominant L2 might be associated with increased rating leniency (e.g., Hansen et al., 2004; Weyant, 2007), which may not be as effective at reducing ratings that have already been enhanced through social priming. Based on the lack of systematic investigation of various rated dimensions in relation to social priming, we had no expectation regarding how task practice might impact listener assessments of L2 accent versus comprehensibility.

Method

Listeners

Listeners included 70 English–French bilinguals (55 females), residents of Montreal, all self-identified members of Quebec's anglophone community. They were recruited through community flyers, social media groups, and word of mouth. Listeners were raised in a family with at least one native English-speaking parent and schooled in Quebec through English- and/or French-medium instruction at the elementary (English = 15, French = 5, English with French immersion = 50), secondary (English = 23, French = 6, English with French immersion = 41), and postsecondary (English = 52, French = 1, English and French = 4) levels. For all listeners, English was self-reported as more dominant than French. Of the 70 listeners, 30 participated in the study by Taylor Reid et al. (2019), providing baseline data for the negative, positive, and no social priming conditions (see below). The 40 listeners tested for this study experienced the same

negative or positive social priming, with half of the listeners also exposed to a task practice intervention in English or French. Listeners' characteristics are summarized in Table 1. One-way ANOVAs comparing background characteristics between the seven groups of listeners, with 10 listeners per group (see Table 1), revealed no significant differences across the groups for listeners' age, $F(6, 63) = 0.42$, $p = .865$, their amount of English use, $F(6, 63) = 1.32$, $p = .260$, their self-rating of English speaking, $F(6, 63) = 1.56$, $p = .174$, their self-rating of English listening, $F(6, 63) = 1.00$, $p = .433$, their amount of French use, $F(6, 63) = 1.76$, $p = .122$, their self-rating of French speaking, $F(6, 63) = 0.91$, $p = .493$, or their self-rating of French listening, $F(6, 63) = 0.87$, $p = .521$. Based on similar one-way ANOVAs, the groups also did not differ in their responses to a social attitudes survey (see Appendix) eliciting their pride for Quebec's anglophones (5 questions, $\alpha = .93$), $F(6, 63) = 0.81$, $p = .569$, their perception of the role of English in Quebec (5 questions, $\alpha = .68$), $F(6, 63) = 0.77$, $p = .598$, their attitudes towards immigrants (5 questions, $\alpha = .77$), $F(6, 63) = 0.94$, $p = .474$, and their feelings towards other ethnolinguistic groups (5 questions, $\alpha = .69$), $F(6, 63) = 1.09$, $p = .377$. All but one listener reported normal hearing (a minor hearing loss), but that individual's data remained, as removing those data resulted in no change in findings.

Table 1

Means (Standard Deviations) for Listeners' Background Characteristics

Social priming intervention	No priming	Positive social priming			Negative social priming		
Task practice	None	None	English	French	None	English	French
Age (years)	25.4 (3.4)	24.7 (3.9)	23.8 (3.0)	23.9 (3.4)	24.8 (3.7)	24.5 (3.7)	23.3 (3.1)
English use (%)	84.5 (18.0)	87.0 (10.6)	91.0 (5.7)	88.0 (7.9)	87.0 (11.6)	81.0 (15.2)	76.0 (20.7)
English speaking (1–9 scale)[a]	9.0 (0.0)	8.8 (0.4)	8.9 (0.3)	9.0 (0.0)	9.0 (0.0)	9.0 (0.0)	9.0 (0.0)
English listening (1–9 scale)[a]	9.0 (0.0)	8.9 (0.3)	9.0 (0.0)	9.0 (0.0)	9.0 (0.0)	9.0 (0.0)	9.0 (0.0)
French use (%)	23.5 (12.9)	31.0 (21.3)	18.0 (14.0)	17.0 (14.2)	17.5 (10.3)	34.0 (19.0)	26.5 (19.7)
French speaking (1–9 scale)[a]	6.1 (1.7)	5.8 (1.2)	5.6 (1.5)	6.1 (1.4)	6.6 (1.5)	6.5 (1.6)	6.8 (1.2)
French listening (1–9 scale)[a]	6.9 (1.3)	6.6 (1.4)	6.5 (1.1)	6.7 (1.6)	7.2 (1.7)	6.7 (1.6)	7.7 (1.3)
Pride in Anglophone group[b]	6.0 (2.7)	7.1 (2.1)	6.7 (2.2)	6.0 (2.7)	7.5 (1.2)	7.4 (1.9)	7.1 (1.9)
Role of English in Quebec[b]	5.6 (1.3)	5.2 (1.2)	6.2 (1.3)	5.7 (1.7)	6.1 (2.0)	5.1 (1.2)	5.7 (0.9)
Attitudes towards immigrants[b]	2.0 (0.7)	2.7 (1.0)	3.1 (1.7)	2.2 (0.9)	2.9 (1.7)	3.1 (1.1)	2.9 (1.7)
Feelings towards other groups[b]	6.8 (1.7)	7.3 (1.1)	6.0 (1.5)	7.5 (1.2)	7.4 (0.9)	7.2 (1.3)	7.3 (1.2)

Note. [a]1 = beginner, 9 = nativelike. [b]Mean of five question responses, based on a 1–9 scale (1 = disagree, 9 = agree); n = 10 per group.

Materials and Procedure

The audio files evaluated by listeners were drawn from the corpus analyzed by Trofimovich and Isaacs (2012) and included recordings made by 40 native speakers of French (27 females) raised in Quebec in French-speaking households and educated exclusively in French. The speakers (age 18–61) narrated a picture story in their L2 (English). The eight-panel story showed two people who collide at a city street corner, accidentally picking up each other's identical suitcases, and later realizing that the contents were not their own belongings (Derwing et al., 2004). Consistent with prior research (Derwing & Munro, 2015), the initial 30 seconds of each narrative was included for rating, minus initial false starts and hesitations.

The audio files were evaluated for accent and comprehensibility using 1,000-point sliding scales that included anchor point descriptors but no numeric or interval markings (for scale validation, see Saito et al., 2017). Accent was defined to listeners as a measure of how different a speaker sounds from a native English speaker (0 = heavily accented, 1,000 = no accent at all). Comprehensibility was explained as how easy or difficult a speaker is to understand (0 = hard to understand, 1,000 = easy to understand). Both scales were introduced using definitions and examples, and listeners could ask for clarification before proceeding to evaluate the 40 audio files in a unique, randomized order. Listeners heard each file once before rating accent and comprehensibility, on the assumption that such ratings reflect initial, intuitive perceptual judgments.

The procedure was identical for all participants except for the pre-rating task assignment (English or French task practice) and social priming (negative, positive). After listeners completed a background questionnaire, they were shown the same picture story described in the audios and were asked to narrate the story in either English (20 listeners) or French (20

listeners). Then, all listeners received rating instructions and completed three sample ratings. At this point, as the researcher was setting up the rating interface, she casually delivered a socially-charged priming anecdote (the same as in Taylor Reid et al., 2019). In the negative priming condition (10 listeners who had practiced the task in English and 10 who had practiced the task in French), the researcher shared a short story about her recent experience at a local café, where she had not been served adequately in English by a native French-speaking employee who had, according to the script, an atrocious accent, poor grammar, and had not bothered to learn English. In the positive priming condition (10 listeners who had practiced the task in English and 10 who had practiced the task in French), the researcher (the same for all sessions) told a comparable story, except that it featured her positive experience at the same café, where she had been served by a native French-speaking employee whose accent and grammar were excellent and who exhibited great effort to learn English (see Appendix for full scripts). The baseline groups from Taylor Reid et al. (2019) did not engage in task practice and completed ratings with negative, positive, or no social priming imposed (10 listeners per group). At the end of the session, listeners completed a debrief questionnaire using 100-millimeter scales targeting their perception of the session, the researcher's helpfulness, the rating task difficulty, and their rating confidence (see Appendix), then filled out a social attitudes survey. Listener comments made at the end of the session were added to the researcher's field notes. The entire session lasted about two hours per participant.

Data Analysis

All ratings, which displayed high consistency within each listener group for accentedness (Cronbach's α = .94–.97) and comprehensibility (Cronbach's α = .91–.96), were transferred into a matrix containing 2,800 observations (70 listeners × 40 speakers) for mixed-effects modeling

(see below). Although there are no agreed-upon criteria for estimating adequate sample sizes for

mixed-effects models (Maas & Hox, 2005; McNeish & Stapleton 2016), the estimates provided

by Scherbaum and Ferreter (2008) suggested that a sample of 30 listeners and 30 speakers is

needed to achieve power of .80 with a medium effect size. Therefore, the current sample of 70

listeners and 40 speakers was deemed sufficiently large for mixed-effects modeling. Debrief

questionnaires were scored as distance (in millimeters) between the left endpoint and listeners'

marks (the intersection of the cross or angle point of the checkmark) on the 100-millimeter scale.

As summarized in Table 2, there were no differences between the listener groups in debrief

ratings, $F(6, 63) < 1.49$, $p > .20$, with all groups finding the experience pleasant (on a *very

unpleasant–very pleasant* scale) and the researcher helpful (on a *not helpful at all–very helpful*

scale), evaluating task difficulty similarly (on a *very difficult–very easy* scale), and being

similarly confident in their ratings (on a *not confident at all–very confident* scale).

Table 2

Means (Standard Deviations) for Listeners' Debrief Ratings

Social priming intervention	No priming	Positive social priming			Negative social priming		
Task practice	None	None	English	French	None	English	French
Experience pleasant	91.60 (10.20)	87.60 (13.85)	89.05 (7.54)	89.95 (7.04)	95.90 (7.82)	87.75 (13.02)	91.85 (7.26)
Researcher helpful	98.20 (2.70)	97.35 (3.48)	96.15 (3.90)	96.20 (5.18)	97.30 (7.20)	95.60 (4.43)	97.55 (2.48)
Task difficult	76.80 (18.74)	68.50 (24.51)	61.90 (23.52)	75.40 (15.48)	71.00 (24.67)	67.35 (23.29)	81.15 (11.57)
Confident in rating	81.45 (9.26)	80.05 (16.88)	61.70 (26.53)	77.35 (17.28)	72.25 (21.64)	69.35 (21.56)	79.05 (8.27)

Listeners assigned to task practice assessed their speaking proficiency higher in English (M = 8.90, SD = 0.16) than in French (M = 6.25, SD = 1.45), $t(39)$ = 12.04, $p < .001$, and also estimated their listening proficiency higher in English (M = 9.00, SD = 0) than in French (M = 6.90, SD = 1.43), $t(39)$ = 9.30, $p < .001$. Participants who narrated the pictured sequence in English took less time (M = 35 seconds, SD = 14) than those who completed it in French (M = 59 seconds, SD = 37), $t(38)$ = 2.62, $p = .013$, which provides further evidence of the listeners' self-reported English dominance. Listeners' debrief comments, analyzed broadly for evidence of their awareness of a priming manipulation, revealed that 8–10 listeners per group (i.e., 80–100%) responded "no" when asked if anything that the researcher said influenced their ratings, with the remaining listeners commenting on several (positive) aspects of their experience (e.g., "[the researcher] was very objective," "she helped me to better understand the proportions of the large scales," "overall I had a great experience"). Thus, listeners seemed unaware of the social priming.

Accent and comprehensibility ratings were analyzed through mixed-effects modeling using the lme4 package (Bates et al., 2015) in R version 4.0.2 (R Core Team, 2020). In each model, priming (none, negative, positive), task practice (no practice, English, French), and their interaction served as fixed factors, and speakers and listeners were entered as random intercepts. By-speaker random slopes for priming and language practice and by-listener random slopes for speakers were included only if they improved model fit, evaluated through likelihood ratio tests on each pair of models using the ANOVA function. To determine the statistical significance of each parameter, 95% confidence intervals (CI) were inspected (interval does not cross zero), and probability values were derived through a Satterthwaite's approximation using the lmerTest package in R.

Results

Table 3 summarizes descriptive statistics for accent and comprehensibility ratings by listener group, where higher ratings on a 1,000-point scale designate less accented (0 = heavily accented, 1,000 = no accent at all) and more comprehensible (0 = hard to understand, 1,000 = easy to understand) L2 speech. Listeners' assessments appeared to vary between groups, as demonstrated through differences in mean accent ratings (*range* = 372–452) and mean comprehensibility ratings (*range* = 510–619); these assessments also varied within each group, as shown through non-trivial standard deviation values. Thus, mixed-effects modeling appeared ideal for examining potential between-group differences in listener behavior while trying to account for within-group individual listener variation in ratings assigned to L2 speakers.

Table 3

Descriptive Statistics for Accent and Comprehensibility Ratings

	Accent		Comprehensibility	
Group	*M*	*SD*	*M*	*SD*
No priming, no task practice	372	168	545	177
Positive social priming				
No task practice	435	231	604	213
English task practice	412	198	581	233
French task practice	425	208	550	235
Negative social priming				
No task practice	452	233	619	225
English task practice	391	215	510	226
French task practice	434	238	578	238

Note. 1000-point rating scale; higher ratings designate less accented and more comprehensible L2 speech.

Accent

The best-fitting model for accent ratings, which accounted for 73% of the variance (conditional R^2 = .73), included by-speaker and by-listener random intercepts, along with by-speaker random slopes for priming and by-listener random slopes for speakers. The final model, summarized in Table 4, significantly improved model fit compared to simpler models that excluded random slopes for speakers, $\chi^2(2) = 6.64$, $p = .036$, or included only by-speaker and by-listener random intercepts, $\chi^2(7) = 64.98$, $p < .0001$. By-speaker random slopes for language practice did not improve model fit, $\chi^2(7) = 9.79$, $p = .200$. As shown in Table 4, priming resulted

in enhanced accent ratings (with speakers rated as being less accented), where listeners upgraded speakers in both the positively-primed (+56 points) and the negatively-primed (+71 points) groups, although the effect reached statistical significance only for negative priming. Nevertheless, consistent with Taylor Reid et al.'s (2019) findings, the impact of priming was most pronounced for those not engaged in language practice (i.e., in comparisons of ±priming groups only), where listeners significantly upgraded L2 speakers' accent under positive priming *Estimate* = 62.59, *t* = 3.52, *p* = .00046, 95% CI [27.67, 97.51] and under negative priming, *Estimate* = 80.00, *t* = 4.44, *p* < .00001, 95% CI [44.64, 115.37].

Table 4

*Summary of Final Mixed-Effects Model for Accent*

Fixed effects	Estimate	*SE*	95% CI	*t*	*p*
(Intercept)	388.53	35.07	[319.84, 457.30]	11.08	< .0001
Social priming					
Positive	55.82	34.69	[–12.15, 123.79]	1.61	.112
Negative	71.41	35.33	[2.18, 140.65]	2.02	.046
Positive priming with task practice					
English	–41.39	33.95	[–107.87, 25.19]	–1.22	.227
French	–24.75	33.95	[–91.24, 41.82]	–0.73	.468
Negative priming with task practice					
English	–74.28	33.96	[–131.79, –9.27]	–2.19	.032
French	–33.85	33.96	[–89.16, 51.90]	–0.99	.322

Random effects	*SD*	Information criteria	Estimate
Listener (intercept)	90.19	Log-likelihood	–17944.90
Speaker (slope)	0.65	AIC	35923.90
Speaker (intercept)	160.96	BIC	36024.80
Positive priming (slope)	44.57		
Negative priming (slope)	61.40		

Note. AIC = Akaike information criterion, BIC = Bayesian information criterion.

Of key interest to this study's goals were the effects of English and French task practice on listener behavior under positive and negative priming, relative to the impact of positive and negative priming alone (in the absence of language practice). As shown in Table 4 and illustrated

in Figure 1, only the negatively-primed listeners who performed the practice task in English (their stronger language) showed a significant decrease in their accent ratings (–74 points, 95% CI [–131.79, –9.27]), compared to those who received negative priming only (*p* = .032), bringing their ratings in line with those of baseline listeners (no priming, no practice).
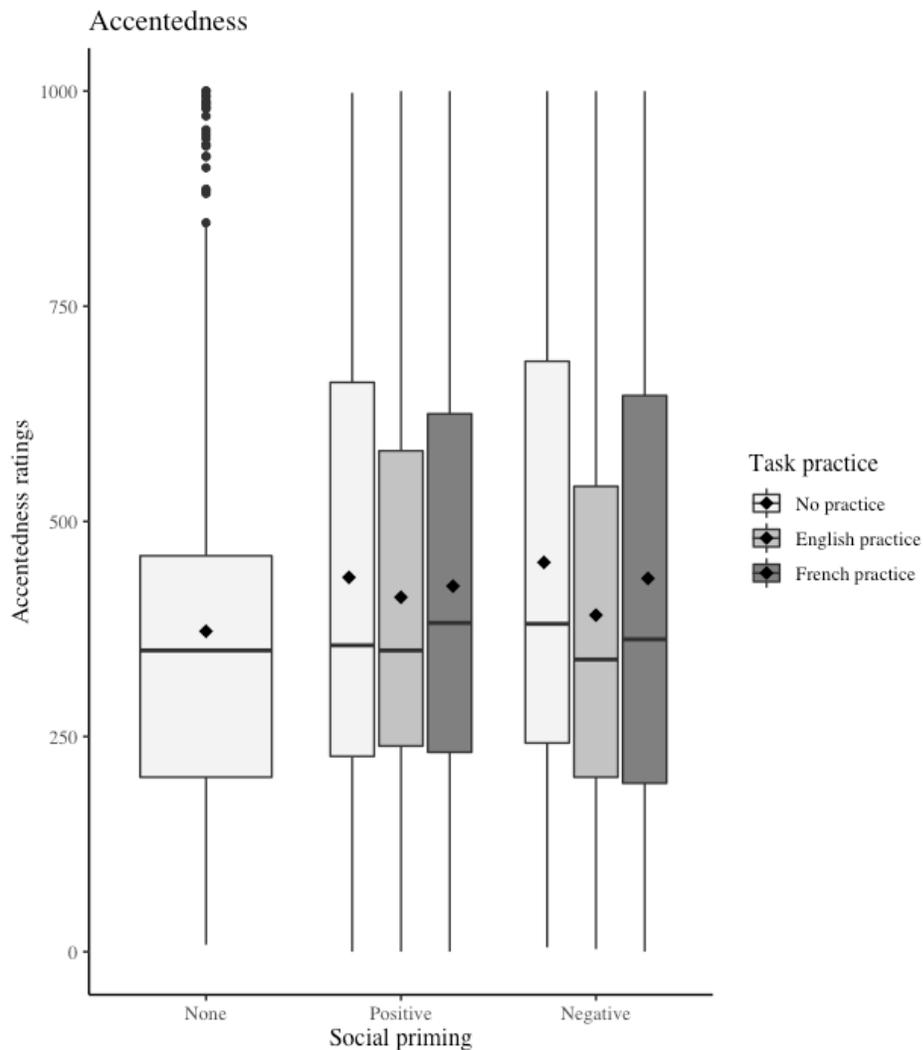
Figure 1. Boxplots for accent ratings (1,000-point scale) by social priming and task practice for a total of 2,800 observations (40 speakers × 70 listeners). Horizontal lines designate median values while diamonds represent means. Higher ratings indicate less accented (more nativelike) L2 speech.

Comprehensibility

The best-fitting model for comprehensibility, which accounted for 68% of the variance (conditional R^2 = .68), included by-speaker and by-listener random intercepts, along with by-speaker random slopes for priming. The final model, summarized in Table 5, significantly improved model fit compared to simpler models that included only by-speaker and by-listener random intercepts, $\chi^2(5)$ = 32.74, p < .0001, or only a by-speaker random intercept, $\chi^2(6)$ = 645.56, p < .0001. By-speaker random slopes for language practice, $\chi^2(9)$ = 5.37, p = .801, or by-listener random slopes for speakers, $\chi^2(7)$ = 0.65, p = .998, did not improve model fit. As shown in Table 5, priming resulted in enhanced comprehensibility ratings (with speakers rated as being more comprehensible), where listeners upgraded L2 speakers in both the positively-primed (+58 points) and the negatively-primed (+74 points) groups, although the general effect did not reach statistical significance across the three primed groups. However, again consistent with Taylor Reid et al.'s (2019) findings, priming was most pronounced for those not engaged in language practice (i.e., in comparisons of ±priming groups), with listeners significantly upgrading speakers' comprehensibility under positive priming, *Estimate* = 58.27, t = 3.10, p = .002, 95% CI [21.38, 95.16], and under negative priming, *Estimate* = 73.83, t = 3.79, p < .0001, 95% CI [35.61, 112.06].

Table 5

Summary of Final Mixed-Effects Model for Comprehensibility

Fixed effects	Estimate	*SE*	95% CI	*t*	*p*
(Intercept)	545.38	41.06	[464.12, 626.64]	13.28	< .0001
Social priming					
Positive	58.27	44.77	[–30.65, 147.19]	1.30	.197
Negative	73.84	44.83	[–15.19, 162.86]	1.65	.103
Positive priming with task practice					
English	–22.96	44.07	[–110.56, 64.64]	–0.52	.604
French	–53.55	44.07	[–141.14, 34.05]	–1.22	.229
Negative priming with task practice					
English	–108.96	44.07	[–196.56, –21.36]	–2.47	.015
French	–41.11	44.07	[–128.70, 46.49]	–0.93	.354

Random effects	*SD*	Information criteria	Estimate
Listener (intercept)	95.10	Log-likelihood	–18431.30
Speaker (intercept)	169.07	AIC	36892.60
Positive priming (slope)	49.79	BIC	36981.70
Positive priming (slope)	51.84		

Note. AIC = Akaike information criterion, BIC = Bayesian information criterion.

     As with accent ratings, central to the goals of this study were the effects of English and French task practice on listener behavior under positive and negative priming, relative to the impact of positive and negative priming alone (in the absence of language practice). As shown in

Table 5 and illustrated in Figure 2, again, only the negatively-primed listeners who performed the practice task in English (their stronger language) showed a significant decrease in their comprehensibility ratings (–109 points, 95% CI [–196.56, –21.36]), compared to those who received negative priming only ($p = .015$), bringing their ratings in line with those of baseline listeners (no priming, no practice).
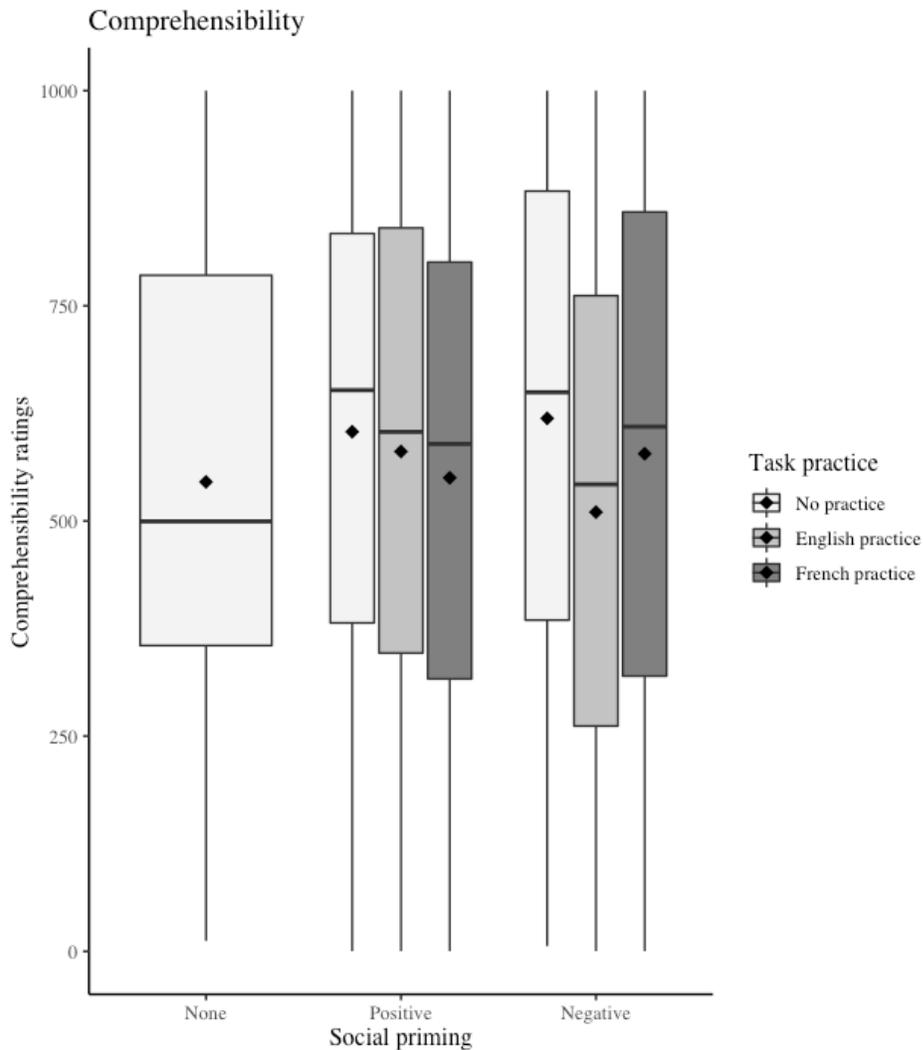

Figure 2. Boxplots for comprehensibility ratings (1,000-point scale) by social priming and task practice for a total of 2,800 observations (40 speakers × 70 listeners). Horizontal lines designate

median values while diamonds represent means. Higher ratings indicate more comprehensible (easier to understand) L2 speech.

Discussion

This study's objective was to determine whether engaging English–French bilingual listeners in task practice could reduce social influences (conceptualized here in terms of social priming) on listener-based assessments of L2 accent and comprehensibility, on the assumption that the effectiveness of task practice depends on whether listeners share a speaking task with L2 speakers (English task practice) or whether they share both a speaking task and a language status with L2 speakers (French task practice). Only English task practice appeared to significantly reduce the effects of priming on listeners' assessments of accent and comprehensibility, and only under negative priming. Findings therefore provide preliminary evidence that task practice might be a medium through which L2 accent and comprehensibility ratings can be stabilized, thereby countering the effects of social priming.

In terms of negative and positive priming effects alone (in the absence of task practice), the young bilingual listeners upgraded their ratings of accent and comprehensibility in both positive and negative priming conditions, relative to baseline assessments. As we argued in Taylor Reid et al. (2019), the young bilingual Montrealers appear to be linguistically and socially aware, showing sensitivity to an anecdote which portrayed an interaction with an L2 speaker in either a positive or a negative light. Listeners accepted the positivity, but also rejected the negativity, demonstrating enhanced solidarity with L2 speakers through their ratings. This solidarity likely reflected listeners' greater exposure to and personal use of L2 speech in the context of their daily experience with English and French in multilingual and multicultural Montreal (Lamarre et al., 2002). Listeners' exposure to the priming anecdotes was clearly not

detrimental to L2 speakers, because listener assessments were enhanced rather than downgraded; nevertheless, this reaction poses a non-trivial threat to the validity and reliability of listener-based assessments of L2 speech, because priming appeared to generate detectable social influence on ratings.

Our main objective was, therefore, to examine task practice as a strategy to reduce social influences on listener-based evaluations of L2 speech. Task practice indeed emerged as having a mitigating effect similar to those shown for the perspective-taking interventions explored by Hansen et al. (2014) and Weyant (2007), but only under specific conditions. Only English task practice appeared to reduce social influences on listener-based assessments of L2 accent and comprehensibility, and only for listeners exposed to negative social priming. For listeners completing task practice in English, it is possible that engaging in a picture narrative in the same language as the L2 speakers may have activated key vocabulary and content and focused attention to narrative detail (e.g., Gass & Varonis, 1984), such that listeners had higher expectations of hearing a narrative that was comparable to their own. Familiar with the content and vocabulary of the task to be evaluated, listeners may thus have been more attuned to the linguistic dimensions of L2 speakers' speech (especially those that contribute to accent and comprehensibility). Performing the task in English was also useful to listeners because it provided them with a task model to follow, given that task familiarity enhances listener calibration (Davis, 2016). Listeners practicing the target task in English may have thus developed a more specific, realistic model of task performance and used their own narratives (e.g., featuring performance errors and native-speaker variations) as the baseline for speech ratings. In essence, increased task familiarity—coupled with realistic performance expectations available to listeners through English task practice—may have limited the impact of the priming

anecdote.

Considering that the young bilingual Montrealers in Taylor Reid et al. (2019) who were exposed to negative social priming upgraded their assessments of L2 speakers, ostensibly "fighting back" against the negativity in the priming anecdote, the stabilizing effect of the English task practice shown here offers a nuanced understanding of the role of task practice in reducing socially-induced variance in speech ratings. The obtained difference in the impact of the English task practice on positive versus negative social priming might be broadly understood within accessibility models of social priming (Strack & Schwarz, 2016), including the situated inference model (Loersch & Payne, 2016). This model assumes that priming effects arise as a function of increased accessibility of information related to the priming stimulus, such that people's currently accessible thoughts and feelings are most prone to the impact of outside influence. In this sense, young bilingual Montrealers' positive reactions to a negative prime (in Taylor Reid et al.) and their opposite (moderating) responses to the English task practice here, which brought them back to the performance level of the baseline listeners, likely involve highly accessible information for these participants. This information presumably includes bilinguals' heightened awareness of multilingual language use in Montreal in a social environment with a history of tension between English- and French-speaking communities (Lamarre, 2013) as well as rising ethnolinguistic frictions between both communities and recent immigrants (Castonguay, 2019; Gravelle, 2017). Unlike the negativity surrounding language use in Montreal, a positively-charged priming anecdote (which simply emphasized a speaker's good language skills) perhaps activated information that was less topical and relevant, and thus relatively less accessible for these listeners (e.g., as if striking a less sensitive chord), with the consequence that priming effects arose, but they were less easy to modulate through task practice. Needless to say, these

nuanced accessibility-based differences in availability of primed information and its susceptibility to task practice effects in bilingual listeners must be revisited in future work.

Task practice in French did not appear to produce a statistically reliable impact on reducing variance in L2 ratings arising through social priming. We had reasoned that listeners tasked with practicing the target narrative in their own L2, prior to engaging in speech ratings, might develop a shared understanding of L2 speakers as a result of stepping in their shoes (Hansen et al., 2004; Weyant, 2007), resulting in leniency towards L2 speech. Thus, at least for the listeners tested here, L2 task practice might not be as effective at reducing ratings that have already been enhanced through social priming (Taylor Reid et al., 2019). Listeners who performed the task in their less comfortable language—in which they reported low to intermediate proficiency—were likely inclined towards solidarity with fellow L2 speakers, which translated into rating generosity. Consistent with this idea, listeners' impromptu comments about their own narrative performances in French were generally self-critical (e.g., "It's hard to think of all the vocabulary!" "I didn't think my French was that bad, but it was so hard to tell the story in my [L2]"), suggesting a heightened awareness of the challenges of using a non-dominant L2. In addition, listeners who completed the French task practice might have experienced foreign language anxiety, which is situation-specific and can result in negative emotions (Dewaele, 2002). Therefore, performing the task in French could have resulted in more steadfast leniency towards fellow L2 speakers than among those who completed the narrative in English, in that those who completed the French narrative fully understood the challenges of doing so in their less-dominant language.

As shown in Figures 1 and 2, however, our interpretations concerning the French task practice must remain tentative at best. The solidarity-based account of the French task practice

implies that listeners might demonstrate additional leniency in their ratings—yet the effects of social priming and of the French task practice were not additive—in that rating generosity brought about by practicing the target task in French did not result in additional increases in ratings, over and above those already shown in response to social priming. This implies that there might be a limit to rating leniency, in the sense that inflating speaker performance too high above a certain threshold would be inconsistent with the specific evidence available to listeners in L2 speakers' speech (e.g., segmental substitutions, pausing frequency). Identifying such potential leniency or severity thresholds—separately for ratings of L2 accent versus comprehensibility—and finetuning explanations for task practice effects in listeners' more versus less dominant languages illustrate interesting challenges to be addressed in future work.

Limitations and Conclusion

Because listeners in this study were bilinguals from a similar background, all residing in a bilingual country and in a context where they belong to the linguistic minority, future research should consider the extent to which each of these factors plays a role in listeners' reactions to L2 speech, particularly because local, contextual, and person-specific influences might impact people's susceptibility to social priming (Loersch & Payne, 2016). Additional work could examine the extent to which listeners' judgments are affected by social priming and can be mitigated in more naturalistic situations (e.g., job interviews) exploring, for instance, L2 speakers' task- or job-relevant competence demonstrated in the speech samples to be rated.

Because the current work was generally driven by practical considerations of mitigating social priming effects on speech ratings, our findings have limited theoretical impact, in that social priming literature was only used here to frame and interpret the study's outcomes. Future studies on this and similar issues should utilize research designs which might allow researchers

to answer theoretically-driven, rather than practice-oriented, questions, such as distinguishing among various explanation of social priming effects. These investigations might include, for instance, further exploration of the individual and combined effects of social priming (which might prepare listeners to hear L2 speech that is more or less accented and comprehensible) and those of task practice (which might enhance listeners' familiarity with and expectations of L2 speech).

Despite these shortcomings, and given the ubiquity of listeners' decision-making in the real world and its serious implications for wages (Dávila et al., 1993), hiring decisions (Hansen & Dovidio, 2016), and court proceedings (Frumkin, 2007), our findings are relevant beyond the laboratory setting. While much of the laboratory-based research has demonstrated consistency among listeners, concluding that human listeners are the gold standard for speech assessment (Eskenazi, 2009), our findings show that listener-based L2 speech ratings may be less stable than previously assumed and that engaging listeners in task practice has the potential to impact rating stability.

References

Baquiran, C. L. C., & Nicoladis, E. (2020). A doctor's foreign accent affects perceptions of competence. *Health Communication*, *35*, 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*, 1–48. https://doi.org/10.18637/jss.v067.i01

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

Campbell-Kibler, K. (2007). Accent, (ING), and the social logic of listener perceptions. *American*

Speech, *82*, 32–64. https://doi.org/10.1215/00031283-2007-002

Castonguay, C. (2019). Quebec's new language dynamic: French fading fast. *Language Problems and Language Planning*, *43*, 113–134.

Dávila, A., Bohara, A., & Saenz, R. (1993). Accent penalties and the earnings of Mexican Americans. *Social Science Quarterly*, *74*, 902–915.

Davis, L. (2016). The influence of training and experience on rater performance in scoring spoken language. *Language Testing*, *33*, 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*, 417–426. https://doi.org/10.1080/00207594.2010.499950

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

Derwing, T. M., Rossiter, M. J., & Munro, M. J. (2002). Teaching native speakers to listen to foreign-accented speech. *Journal of Multilingual and Multicultural Development*, *23*, 245–259. https://doi.org/10.1080/01434630208666468

Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thomson, R. I. (2004). Second language fluency: Judgments on different tasks. *Language Learning*, *54*, 655–679. https://doi.org/10.1111/j.1467-9922.2004.00282.x

Dewaele, J. M. (2002). Psychological and sociodemographic correlates of communicative anxiety in L2 and L3 production. *International Journal of Bilingualism*, *6*, 23–38. https://doi.org/10.1177/13670069020060010201

Dijksterhuis, A., & Bargh, J. A. (2001). The perception-behavior expressway: Automatic effects

of social perception on social behavior. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (pp. 1–40). Academic Press. https://doi.org/10.1016/S0065-2601(01)80003-4

Eskenazi, M. (2009). An overview of spoken language technology for education. *Speech Communication*, *51*, 832–844. https://doi.org/10.1016/j.specom.2009.04.005

Frumkin, L. (2007). Influences of accent and ethnic background on perceptions of eyewitness testimony. *Psychology, Crime & Law*, *13*, 317–331. https://doi.org/10.1080/10683160600822246

Gaither, S. E., Cohen-Goldberg, A. M., Gidney, C. L., & Maddox, K. B. (2015). Sounding Black or White: Priming identity and biracial speech. *Frontiers in Psychology*, *6*, 457. https://doi.org/10.3389/fpsyg.2015.00457

Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: Decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology*, *78*, 708-724. https://doi.org/10.1037/0022-3514.78.4.708

Gass, S. M., & Varonis, E. M. (1984). Input, interaction, and second language production. *Studies in Second Language Acquisition*, *16*, 283–302. https://doi.org/10.1017/S0272263100013097

Gatbonton, E., & Trofimovich, P. (2008). The ethnic group affiliation and L2 proficiency link: Empirical evidence. *Language Awareness*, *17*, 229–248. https://doi.org/10.1080/09658410802146867

Gravelle, T. B. (2017). Partisanship, local context, group threat, and Canadian attitudes towards immigration and refugee policy. *Migration Studies*, *6*, 448–467. https://doi.org/10.1093/migration/mnx058

Haladyna, T. M., & Downing, S. M. (2004). Construct-irrelevant variance in high-stakes testing. *Educational Measurement: Issues and Practice*, *23*, 17–27. https://doi.org/10.1111/j.1745-3992.2004.tb00149.x

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*, 68–78. https://doi.org/10.1177/1359105315595121

Hansen, K., & Dovidio, J. F. (2016). Social dominance orientation, nonnative accents, and hiring recommendations. *Cultural Diversity and Ethnic Minority Psychology*, *22*, 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*, 68–77. https://doi.org/10.1177/0261927X13499761

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*, 441–456. https://doi.org/10.1177/0261927X09341950

Kerr, N. L., Niedermeier, K. E., & Kaplan, M. F. (1999). Bias in jurors vs bias in juries: New evidence from the SDS perspective. *Organizational Behavior and Human Decision Processes*, *80*, 70–86. https://doi.org/10.1006/obhd.1999.2855

Lamarre, P. (2013). Catching "Montréal on the Move" and challenging the discourse of unilingualism in Québec. *Anthropologica*, *55*, 41–56.

Lamarre, P., Paquette, J., Kahn, E., & Ambrosi, S. (2002). Multilingual Montreal: Listening in on the language practices of young Montrealers. *Canadian Ethnic Studies*, *34*, 47–78.

Loersch, C., & Payne, B. K. (2016). Demystifying priming. *Current Opinion in Psychology*, *12*,

32–36. https://doi.org/10.1016/j.copsyc.2016.04.020

Maas, C. J., & Hox, J. J. (2005). Sufficient sample size for multilevel modeling. *Methodology*, *1*, 86–92. https://doi.org/10.1027/1614-1881.1.3.86

McNeish, D. M., & Stapleton, L. M. (2016). The effect of small sample size on two-level model estimates: A review and illustration. *Educational Psychology Review*, *28*, 295–314. https://doi.org/10.1007/s10648-014-9287-x

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

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*, 222–243.

Plonsky, L., & Derrick, D. J. (2016). A meta-analysis of reliability coefficients in second language research. *The Modern Language Journal*, *100*, 538–553. https://doi.org/10.1111/modl.12335

R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Saito, K., Trofimovich, P., & Isaacs, T. (2017). Using listener judgements to investigate linguistic influences on L2 comprehensibility and accentedness: A validation and generalization study. *Applied Linguistics*, *38*, 439–462. https://doi.org/10.1093/applin/amv047

Scherbaum, C. A., & Ferreter, J. M. (2009). Estimating statistical power and required sample

    sizes for organizational research using multilevel modeling. *Organizational Research*

    Methods, *12*, 347–367. https://doi.org/10.1177/1094428107308906

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

Strack, F., & Schwarz, N. (2016). Editorial overview: Social priming–information accessibility and

    its consequences. *Current Opinion in Psychology*, *12*, 58–62.

    https://doi.org/10.1016/j.copsyc.2016.11.001

Taylor Reid, K., Trofimovich, P., & O'Brien, M. (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*, 419–442.

    https://doi.org/10.1017/S0272263118000244

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*, 409–428. https://doi.org/10.1177/0950017016630260

Trofimovich, P., & Isaacs, T. (2012). Disentangling accent from comprehensibility.

    Bilingualism: Language and Cognition, *15*, 905–916.

    https://doi.org/10.1017/S1366728912000168

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*, 703–716. https://doi.org/10.1111/j.1559-1816.2007.00181.x

Winke, P., Gass, S., & Myford, C. (2013). Raters' L2 background as a potential source of bias in

rating oral performance. *Language Testing*, *30*, 231–252.

https://doi.org/10.1177/0265532212456968

Appendix

Social Attitudes Questionnaire

Indicate the degree to which each of these statements accurately reflects how you feel.

	Disagree								Agree

PRIDE FOR ETHNIC GROUP

1. I am proud to be a member of my ethnic group.   1 2 3 4 5 6 7 8 9
2. I am proud to let people know that I belong to my ethnic group.   1 2 3 4 5 6 7 8 9
3. I am proud of the achievements of my ethnic group.   1 2 3 4 5 6 7 8 9
4. I feel proud to see symbols of my ethnic group (such as a flag) displayed around me.   1 2 3 4 5 6 7 8 9
5. I am proud to be able to speak the language of my ethnic group.   1 2 3 4 5 6 7 8 9

ENGLISH IN QUEBEC

6. Anglophone Quebecers do not have considerable economic power in Quebec.   1 2 3 4 5 6 7 8 9

7. Anglophone Quebecers do not have considerable political power in Quebec.   1 2 3 4 5 6 7 8 9

8. In my daily life (for example, in a restaurant, shop, doctor's office), I should have the right to speak English in Quebec.   1 2 3 4 5 6 7 8 9

9. I should have the freedom to choose if I want my children to be educated in English in Quebec.   1 2 3 4 5 6 7 8 9

10. Anglophone Quebecers' contribution to Quebec is not recognized or valued.   1 2 3 4 5 6 7 8 9

ATTITUDES TOWARDS IMMIGRANTS

11. The influx of immigrants is lowering the standard of living of people in Quebec.   1 2 3 4 5 6 7 8 9

12. Laws, customs, and traditions that are specific to immigrant groups should not be imposed on the Quebec society as a whole.   1 2 3 4 5 6 7 8 9

13. Immigrants should adopt the Quebec way of life and values to replace their traditional way or life and values.   1 2 3 4 5 6 7 8 9

14. Immigrants are bringing conflicts in their home countries into Quebec.   1 2 3 4 5 6 7 8 9

15. Immigrants benefit a lot from being in Quebec so they should be loyal to Québec.   1 2 3 4 5 6 7 8 9

PERSONAL RELATIONS WITH OTHER GROUPS

16. I feel accepted and respected by other ethnic groups in Quebec.   1 2 3 4 5 6 7 8 9

17. Members of other ethnic groups do not mind me living in close   1 2 3 4 5 6 7 8 9
proximity to them.

18. Members of other ethnic groups would not object to my children   1 2 3 4 5 6 7 8 9
marrying their children.

19. Children who grow up in an ethnically diverse Quebec are more prepared 1 2 3 4 5 6 7 8 9
to live in today's world.

20. Children growing up in an ethnically diverse Quebec are more tolerant of 1 2 3 4 5 6 7 8 9
other groups.

Note. Category labels were not presented to participants. Questionnaire is based on materials from Gatbonton and Trofimovich (2008).

Scripts for Bias Manipulation

Negative
I'm sorry. I just have to vent about something. You're a native English speaker, so you'll get where I'm coming from. I just went to grab some food from the Copper Branch down the street. I can't believe they don't provide adequate service in English! I mean, the person I placed my order with was clearly a French speaker who had never bothered to learn English very well. I could barely understand her! Her accent was awful and her grammar didn't even make sense. You know – Canada has *two* official languages, so everyone should be able to speak English—at school, at work, everywhere! I can't believe there are French speakers who haven't bothered to become fluent in English after spending a lifetime in Canada! It's ridiculous.

Positive
I just have to share something with you that I've been thinking about all day. You're a native English speaker who probably speaks some French, so you'll get where I'm coming from. So, I went to the Copper Branch down the street earlier to get something to eat and I was so impressed that they provide such good service in English! The person I placed my order with was obviously a native speaker of French, but her English was *really* good! I mean, she had an accent, but it still sounded good and her grammar wasn't bad, either. You know – Canada has *two* official languages, so everyone should be able to use English—at school, at work, everywhere! I am just so impressed that some French speakers make such an effort to learn English. It's tough learning a new language!

Final Debrief Questionnaire

Please rate your experience in today's session by putting an X in the appropriate spot on the scale.

1. How pleasant was your experience in this rating session?

Very unpleasant   Very pleasant

2. How helpful was the researcher during the session?

Not helpful at all   Very helpful

3. How difficult was the rating task for you?

   Very difficult   Very easy

4. How confident are you in your ratings?

Not confident at all   Fully confident

5. Did any part of your interaction with the researcher affect your ratings?

_____

_____

_____

_____