Applied Linguistics, 2024, XX, 1–20
https://doi.org/10.1093/applin/amae042
Advance access publication 11 July 2024
Article

Accent Bias in Professional Evaluations: A Conceptual Replication Study in Brazil

Cesar Teló[1,*], Rosane Silveira[2], Ana Flávia Boeing Marcelino[3], and Mary G. O'Brien[1,4]

[1]School of Languages, Linguistics, Literatures and Cultures, University of Calgary, Canada
[2]Department of Foreign Languages and Literatures, Federal University of Santa Catarina, Brazil
[3]Department of Romance Languages and Literatures, University of Chicago, USA
[4]Department of Linguistics, Simon Fraser University, Canada

[*]E-mail: cesar.telo@ucalgary.ca

Evidence from Canada suggests that accent bias can be moderated by speakers' demonstrated job-relevant performance and the prestige level of their occupation (Teló *et al.* 2022). In this study, we replicated Teló *et al.*'s (2022) work in Brazil. First language (L1) Brazilian Portuguese-speaking listeners rated audio recordings of L1 Brazilian Portuguese and L1 Spanish speakers along continua capturing one professional (competence), one experiential (treatment preference), and one linguistic (comprehensibility) dimension. Our findings challenge the notion of consistent bias, as listeners did not uniformly perceive L1 Brazilian Portuguese speakers as more competent and comprehensible than L1 Spanish speakers, and, in fact, generally preferred treatment provided by L1 Spanish speakers. Complex interactions provided a nuanced account of listeners' evaluations, revealing, among other patterns, that demonstrated performance level and job prestige affected the evaluated dimensions differently depending on the speaker's L1. This replication further expands the initial study by examining the role of four listener variables as predictors of speaker ratings. Greater listener familiarity with the context depicted in the script was associated with the assignment of higher ratings overall.

Introduction

Accent bias—the unfair treatment one receives based on their accent—is pervasive. Second language (L2) speakers whose L2 accent is influenced by their previously known languages (henceforth referred to as L2-accented speakers) face prejudice and discrimination, including in professional settings (Lippi-Green 2012). While the sources of negative language attitudes can be multiple, including intergroup processes such as categorization and stereotyping (Ryan 1983) and fluency-based factors such as processing fluency (comprehensibility; Dragojevic and Giles 2016), they result in negative consequences for L2-accented speakers. For example, they are frequently perceived as less intelligent, trustworthy, credible, and competent than first language (L1) speakers (e.g. Lindemann 2003; Lev-Ari and Keysar 2010; Baquiran and Nicoladis 2020). These biased evaluations can substantially impact the professional integration of L2-accented speakers. In fact, recent meta-analytic evidence (Spence *et al.* 2024) shows that L2-accented speakers are systematically perceived as less employable than L1-speaking

© The Author(s) (2024). Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae042/7712341 by University of Chicago D'Angelo Law Library user on 13 July 2024

candidates. Adding nuance to this pattern, research further suggests that occupational characteristics can affect professional-related evaluations assigned to L2-accented speakers, who are often preferred for low-prestige positions (e.g. Kalin and Rayko 1978; Singer and Eder 1989) and those with lower communication demands (e.g. Lindberg and Trofimovich 2023; Spence *et al.* 2024).

Until recently, however, it was unclear whether and how listeners consider a speaker's job performance or skill level when evaluating the speaker, such that accent bias may interact with speakers' demonstrated job-relevant performance/skill level. To address this gap, Teló *et al.* (2022) examined listeners' evaluations of L1 and L2 English speakers assuming the role of professionals performing more or less skillfully in jobs that also differed in prestige. English-speaking listeners from Calgary (Canada) rated recordings of L1 English and Tagalog speakers along three continua capturing one professional (competence), one experiential (treatment preference; i.e. the extent to which the listener wished to be treated as described in the script), and one linguistic (comprehensibility, i.e. ease of understanding) dimension. Data were analyzed with generalized linear mixed-effects models controlling for the number of languages known by the listeners and their familiarity with L2-accented speech. Overall, L1 English speakers were rated higher than L1 Tagalog speakers on all dimensions. Predictably, high-prestige jobs (doctor, lawyer, professor) and high-performance scenarios tended to be evaluated more favorably than low-prestige jobs (cleaner, salesperson, server) and low-performance scenarios. Nonetheless, three-way interactions among speaker L1, job prestige, and performance level revealed a nuanced picture of accent bias against L2-accented speakers, who were perceived just as competent as and received similar preference ratings to L1 English speakers when job prestige and/or professional performance was low. These findings suggest that listeners may have assumed that Tagalog speakers were particularly suited for low-paying, low-prestige jobs, and that poor professional performance was associated with speakers in such positions.

Notwithstanding their contribution, Teló *et al.*'s (2022) findings are likely contextual. The authors propose, for example, that it is probable that listeners' ratings were influenced by their experience with L2-accented speakers in low-paying, low-prestige positions, or by their expectations regarding the jobs and skills usually associated with those speakers. This raises the question of whether similar findings would be obtained in a country with a different linguistic landscape and a smaller immigration influx, which may affect people's experience with newcomers and the power relations between local-born, L1-speaking individuals and foreign-born, L2-speaking individuals. This motivated the current replication study in Brazil.

While Canada welcomed more than 400,000 new permanent residents in 2021 (representing over 1.08% of its population), just over 150,000 people immigrated to Brazil in the same period, constituting 0.07% of its population. Between 2011 and 2021, the Brazilian job market witnessed a substantial increase in the number of immigrants formally employed, with most of them residing in southern states and working in low-paying occupations (Cavalcanti *et al.* 2022). Recently, Venezuela, Haiti, and Cuba have led immigration, with most employed immigrant women in 2021 coming from Spanish-speaking countries (Cavalcanti *et al.* 2022). Similarly to other contexts, immigrants in Brazil contend not only with broader forms of discrimination but also grapple with linguistic prejudices (Bueno 2018; Mundin and Santos 2022), where the commonly held view of linguistic unity and homogeneity (Bagno 1999) may exacerbate stereotypes and misconceptions about the proficiency and intelligibility of L2-accented speakers (Hamel 2013; Rodrigues 2014).

The present study aims to conceptually replicate and extend Teló *et al.*'s (2022) study in Brazil, an under-investigated context in accent bias research. Because most recent immigrants in Brazil are from Spanish-speaking countries and settle in southern states (Cavalcanti *et al.* 2022), we decided to investigate how L1 Brazilian Portuguese (BP) listeners, residents of the State of Santa Catarina, in the south of Brazil, evaluate L2 BP speakers with L1 Spanish in workplace-relevant contexts. We address the following research question: How are listeners' professional (competence), experiential (treatment preference), and linguistic (comprehensibility) reactions to speakers in job scenarios associated with speakers' linguistic background (Brazilian Portuguese vs.

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae042/7712341 by University of Chicago D'Angelo Law Library user on 13 July 2024

Spanish), job prestige (high vs. low), and performance level (high vs. low)? Beyond the two control covariates originally analyzed (number of languages known, L2 accent familiarity), we included two additional control covariates in our models: listeners' familiarity with the context of each script and listeners' ethnic pride. First, Teló *et al.* (2022) attribute some of their findings to the possibility that listeners' experience with L1 and L2 speakers in certain positions influenced their evaluations. Second, Santa Catarina is known for leaning toward right-wing ideologies (Singer 2021), and was predominantly settled by European immigrants, primarily of German and Italian descent (e.g. Seyferth 2011), with numerous studies highlighting the strong ethnic pride and cultural identity maintained by these communities (e.g. Rinke 2014). Therefore, by incorporating these control covariates, we aim to offer a more comprehensive understanding of the factors possibly influencing speaker evaluations.

Based on previous research and the sociocultural context in Santa Catarina, we expected L1 Spanish speakers to receive more negative evaluations compared to L1 BP speakers. Moreover, building upon Teló *et al.*'s (2022) findings, we hypothesized that speakers in high-prestige jobs and those demonstrating high professional performance would generally receive more favorable evaluations. Finally, we posited that listeners' ethnic pride may act as a moderating factor in L2 accent bias, such that individuals with higher levels of ethnic pride may exhibit accent bias to a greater extent, as they may hold stronger preconceived notions about the suitability of certain linguistic backgrounds for particular job roles. We have labeled this study a conceptual replication given our 'intentional adaptation of the initial study to investigate generalizability to new conditions' (Marsden *et al.* 2018: 235).

Method

All the Brazilian Portuguese materials, except for the recordings, as well as the data that informed our findings, are available on https://osf.io/z3uak/.

Recordings

The recordings were collected by presenting speakers with 12 scripted scenarios narrated by a hypothetical professional. Half of the scenarios depicted low-prestige jobs (cleaner, salesperson, server), while the other half depicted high-prestige jobs (doctor, lawyer, professor). Although occupational prestige may vary between Canada and Brazil, we used the same jobs as in the original study because they presumably are still low- versus high-prestige professions in Brazil. Moreover, the low-prestige occupations included are closely related to the most common occupations held by immigrant women in Brazil, which are cleaner, butcher, salesperson, and other professions related to the industry and food chains (Cavalcanti *et al.* 2022). Within each profession, one script represented high-level performance, while the other portrayed low-level performance by the professional. For example, the high-performance doctor script featured a phone message to a patient. The doctor explained the results of the patient's lab tests in a detailed manner using layperson-friendly language and informed them that a prescription was ready due to the indication of an infection. Conversely, the low-performance doctor script provided similar information but with medical jargon and uncertainty regarding the test results. The terms 'high-level performance' and 'low-level performance' used here broadly refer to better and worse professional performance, respectively.

The scripts used were translations and adaptations of those originally employed. Adaptions considered the situations and language that residents of Santa Catarina might encounter in their interactions with the selected professionals. For example, while in the original study the speaker took on the role of a client relations professional when reading the scripts associated with a cleaner, the professional depicted in the adapted script was the cleaner herself. This adaptation assumed that the interactions between customers and cleaners are not typically mediated by a third party in Brazil. Regarding linguistic differences, because cross-cultural differences may generate different language use in terms of assertiveness and politeness, for example (Thomas,

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae042/7712341 by University of Chicago D'Angelo Law Library user on 13 July 2024

1983; Mendes de Oliveira, 2015), a noteworthy adaptation concerns the use of less 'direct' or 'blunt' language, as judged by the researchers, especially in the scripts illustrating low-level performance. To ensure that the scripts portrayed different performance levels, four L1 BP residents of Santa Catarina assessed the scripts' overall quality and provided comments. Based on this initial feedback, the scripts were revised and subsequently evaluated by 13 additional L1 BP raters, who used 9-point scales (1 = *not at all*, 9 = *very much*) to assess the person depicted in the script for pleasantness, job effectiveness, and sensitivity to the needs of others. Compared to low-performance scripts, high-performance scripts illustrated a professional who was more pleasant (M = 8.47 vs. 5.71), $t(154) = 9.79$, $p < .001$, effective (M = 8.44 vs. 5.24), $t(154) = 11.6$, $p < .001$, and sensitive to the needs of others (M = 8.11 vs. 5.05), $t(154) = 10.07$, $p < .001$.

The final version of the scripts, which had 107.5 words on average (SD = 8.23), was recorded by six individuals residing in Santa Catarina, three L1 BP and three L1 Spanish speakers, using their phones. We included only female speakers to simplify the experimental design (see below).[1] The speakers were recruited through word of mouth and instructed to familiarize themselves with the passages. They were directed to read each script naturally and to speak at a normal pace, with the opportunity to record each passage multiple times. The final selection of the recordings included those that had the most suitable pace and most natural flow as evaluated by the researchers. The speakers (M_{age} = 39.6, SD = 15.6) were all residents of Santa Catarina. The L1 BP speakers were either born and raised in the state or lived most of their lives there. The Spanish speakers were from different Latin American countries (Chile, Colombia, Cuba). The phonological inventories of BP and Latin American Spanish (LAS) differ remarkably in terms of vowels (e.g. BP has more vowels, including an open/closed vowel contrast for mid-front and mid-back oral vowels, and nasal vowels). Conversely, consonants are similar overall, with a few exceptions (e.g. the BP rhotic is mostly a fricative, while the LAS rhotic is a trill; BP has a voiced/voiceless contrast for alveolar fricatives, while LAS has only the voiceless counterpart; LAS has a voiceless affricate, which is an allophone in BP; BP has two palatal fricatives). Despite these and other phonotactic differences and allophonic variations in BP and LAS, the three varieties of Spanish used in this study should not result in noticeable phonological differences among the productions by the L1 Spanish speakers.

Finally, we examined the ratings of accentedness provided by the 60 listeners as part of the experimental procedure (see below). The ratings assigned to the Spanish speakers ranged between 1 and 97 (M = 40.87) on a 100-point scale, where 100 represented no accent at all. The ratings given to the BP speakers ranged between 20 and 100 (M = 91.27), with a reliable difference between the two sets, $t(358) = 25.85$, $p < .001$, M_{diff} = 50.41, 95% CI [46.57, 54.21], d = 18.51. Therefore, to the extent that accentedness ratings capture variations in pronunciation (at the levels of segments and prosody) and voice quality (Derwing and Munro 2015), the recordings by L1 and L2 BP speakers appeared to demonstrate distinct pronunciation performances.

Listeners

To evaluate the audios, 60 listeners (M_{age} = 20.90, SD = 9.89, min = 18, max = 65) were recruited through social media and word of mouth, with inclusion criteria being age (18 or over) and residence (Santa Catarina). All listeners had L1 BP and hailed from 16 cities across the state. Listeners reported using BP in speaking 94.82 percent of the time, on average (SD = 11.67). The largest group of listeners (N = 17, 28%) reported knowing three languages; 36 listeners reported knowing one, two, or four languages, with 12 listeners (20%) in each of these categories; seven listeners (12%) reported knowing five languages. They self-rated their familiarity with L2-accented BP with a mean score of 4.97 (SD = 2.55) on a scale ranging from 1 (*not familiar at all*) to 9 (*very familiar*). Listeners also provided information about their ethnicity through an open-ended question. Most listeners (N = 26, 43%) indicated 'European' as their ethnic group; 6 listeners (10%) indicated 'White' or 'Caucasian' as their ethnicity. The remaining listeners who answered this question reported affiliation to the following ethnic groups: 'Latin American' (N = 5), 'Brazilian' (N = 5), 'Brown' or 'mixed race' (N = 3), 'Indigenous/European' (N = 2), 'Indigenous' (N = 1), 'African' (N = 1),

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae042/7712341 by University of Chicago D'Angelo Law Library user on 13 July 2024

and 'Asian' ($N = 1$). All but two listeners reported having normal hearing; because excluding those listeners' data resulted in no change in the findings, the entire listener sample was considered for analysis. In contrast to Teló *et al.* (2022), who compensated their listeners with CAD$30, the listeners in the present replication received no compensation, in accordance with the research ethics guidelines in Brazil.

Materials and procedures

Listeners evaluated the recordings remotely, using a *Qualtrics* web-based interface. The interface included audio recordings followed by several 100-point sliding scales containing endpoint descriptors only. Three scales aimed to capture listeners' professional, experiential, and linguistic reactions to speakers. The professional scale assessed the perceived competence of the speaker (This person seems competent… *not at all – very much*). The experiential scale gauged the extent to which listeners desired to be treated as described in the script (I would want to be treated as described in this script… *not at all – very much*). The linguistic scale focused on the speaker's comprehensibility (This person is… *hard to understand – easy to understand*). An additional scale targeted speakers' accentedness, to capture between-speaker variation in pronunciation (This person is … *heavily accented – not accented at all*). Comprehensibility and accentedness were defined for listeners as in the original study (the degree of effort it takes to understand the speaker and the degree to which the pronunciation deviates from the expected native-speaker variety, respectively). The remaining ratings were presented without a specific definition but were preceded by a contextualizing sentence (e.g. This person seems competent).

The final set of recordings comprised 24 files, with each of the six speakers contributing audios for one high-prestige job (low- and high-level performance) and one low-prestige job (low- and high-level performance), totaling four audios per speaker. To ensure that no listener evaluated high- and low-performance scenarios for the same job, heard the same speaker's voice twice, or heard an unbalanced number of L1 versus L2 speaker recordings, we organized the 24 audios (6 speakers × 4 scripts) in four balanced experimental lists containing six scripts each. Across all lists, listeners heard each speaker in a high- and a low-prestige job and illustrating high- and low-level performance. However, within each list, listeners experienced a subset of the recordings that ensured equal distribution of speaker L1 (3 BP, 3 Spanish), occupations (6 jobs), and performance levels (3 high-, 3 low-performance scenarios). As the present replication only includes female speakers, we were able to streamline the experimental design, such that the 60 listeners were randomly assigned to one of the four versions of the interface, with numerically more listeners (15) rating each set of audios compared to the initial study.

Prior to completing the rating task, listeners filled out a consent form and a language background questionnaire. Listeners then completed a practice block where they evaluated three L2 BP speakers of different L1 backgrounds (not Spanish) for accentedness and comprehensibility, receiving brief feedback following each evaluation to help listeners contemplate the difference between the concepts (e.g. *We believe you made considerable effort to understand the message. We expect your rating to fall in the first half or quarter of the scale*). Although Teló *et al.* (2022) acknowledge that such feedback may impact listeners' ratings, we opted for including it to ensure that listeners distinguished between accentedness and comprehensibility. In the main rating task, which included six scenarios presented in randomized order, listeners first read a brief description contextualizing each situation (e.g. *You have just phoned an electronics store to ask about the best tablet for your child. The person who received your call left the following voice message for you*). Listeners then clicked on the recording to initiate playback. After hearing the audio for the first time, they rated the speaker for comprehensibility and accentedness. They then listened to the audio a second time and provided competence and treatment preference evaluations. Listeners had the opportunity to provide comments explaining their ratings. No repeated playback was allowed after the second listening.

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae042/7712341 by University of Chicago D'Angelo Law Library user on 13 July 2024

Data analysis

The analyses from Teló *et al.* (2022) were replicated almost identically, except for the inclusion of two new control covariates. To ensure internal consistency, all ratings were first checked using two-way, consistency, average-measure intraclass correlations (ICCs) calculated with the psych package (version 2.3.6; Revelle 2023) in R (version 4.3.0; R Core Team 2023). The ICC values for the full dataset indicated high consistency for the ratings of competence (.94), treatment preference (.95), and comprehensibility (.97).

Generalized linear mixed-effects models were then fit using the lme4 package (version 1.1.33; Bates *et al.* 2015). To accommodate the nature of the data, which consisted of integers with fixed upper and lower bounds, the target ratings were transformed into proportions using min-max normalization. A binomial distribution with a logit link function was then used in the modeling of the data, which is an appropriate choice when the response variable is a proportion (Baum 2008). For each model, we conducted 100,000 iterations using the BOBYQA optimizer, with the number of adaptive Gauss-Hermite quadrature points (nAGQ) set to zero. Listener and speaker served as random intercepts. The following variables were used as fixed-effects predictors: (i) speaker L1 (BP or Spanish), (ii) accentedness ratings (continuous, included to capture between-speaker variation in L2 accent), (iii) job prestige (high or low), (iv) performance level (high or low), (v) interaction term between speaker L1 and accentedness rating, (vi) interaction term between speaker L1 and job prestige, (vii) interaction term between speaker L1 and performance level, and (viii) three-way interaction term among speaker L1, job prestige, and performance level. Categorical variables were dummy-coded.

Following the original study, all models included the number of languages listeners reported knowing (1–5 in our dataset) and their familiarity with L2-accented speech (1–9 scalar rating) as control covariates in their fixed-effects structure.[2] Because Teló *et al.* (2022) highlight the possibility that listeners' experience with L1 and L2 speakers in certain positions affects their evaluations, we included a control covariate capturing listeners' familiarity with the context depicted in each script. This rating was collected as part of the experimental procedure (How familiar are you with the context described in this script … *not at all – very much*), and indicted that listeners were relatively familiar with the contexts (M = 69.75, SD = 28.27). Furthermore, in light of the raciolinguistic component of accent, we included a control covariate capturing listeners' ethnic pride. Listeners' ethnic pride was estimated with five positively-oriented items (e.g., *I am proud to be a member of my ethnic group*; *I am proud of the achievements of my ethnic group*) adapted from Taylor Reid *et al.* (2019) and translated into BP. The five indicator items were deemed reliable (Cronbach's alpha = .88). On average, listeners reported ethnic pride levels around the mid-point of the scale (M = 5.86, SD = 2.25). Correlations between the response variables and the control covariates are available in Supplementary Table 1 of the Supplementary Data and Information.

Despite having 360 observations for each response variable, the lme4 function used row-wise deletion to handle missing values. In the end, there were 351 observations for each rating. To assess the statistical significance of each parameter, we examined the p-values calculated by the lmerTest package (version 3.1.3; Kuznetsova *et al.* 2017) and inspected bootstrapped 95% confidence intervals (CIs). To explore significant interaction effects, post hoc comparisons were conducted using the emmeans package (version 1.8.7; Lenth 2023), with p-values adjusted using the Tukey method.

Results

The research question aimed to investigate how listeners' professional (competence), experiential (treatment preference), and linguistic (comprehensibility) evaluations are associated with speakers' L1 (BP vs. Spanish), job prestige (high vs. low), and performance level (high vs. low). Additionally, the analysis accounted for between-speaker variation in pronunciation by considering the accentedness ratings. The target variables are summarized descriptively in Supplementary Tables 2 and 3 of the Supplementary Data and Information.

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae042/7712341 by University of Chicago D'Angelo Law Library user on 13 July 2024

Competence

As summarized in Table 1, for the ratings of speaker competence, there were significant main effects for accentedness and performance level. Listeners generally evaluated speakers who were rated as having a weaker accent as more competent than those rated as having a stronger accent. Listeners also tended to rate all speakers (regardless of their L1) as more competent in high-level performance scenarios than in low-level performance scenarios. While speaker L1 per se was not significant, the effect for speaker L1 was qualified by an interaction with performance and by an interaction with prestige, signifying that specific combinations of speaker L1 and performance/prestige level affected listeners' evaluations in unique ways. Nonetheless, due to the involvement of these predictors in a statistically significant three-way interaction among speaker L1, performance level, and job prestige (which indicates that all three variables interacted with each other to influence competence ratings; see full comparison details below), the interpretation of main effects and two-way interactions may be misleading. The only control covariate that emerged as significant was context familiarity, indicating that stronger context familiarity was associated with higher competence ratings.

The three-way interaction among speaker L1, performance, and prestige (Figure 1) does not reveal any significant differences in competence ratings when L1 BP and Spanish speakers are compared directly across scenarios. Nonetheless, while L1 BP speakers received comparable competence ratings in high- and low-prestige scenarios at a given performance level, Spanish speakers were rated more competent in low- than high-prestige jobs if their performance level was

Table 1: Summary of mixed-effects model for speaker competence

Fixed effects	**Estimate**	***SE***	**95% CI**	***z***	***p***
(Intercept)	0.56	0.62	[–0.71, 1.67]	0.91	.365
Speaker L1	–0.15	0.29	[–0.64, 0.45]	–0.51	.607
Accentedness	1.09	0.17	[0.77, 1.43]	6.46	<.001
Prestige	–0.06	0.09	[–0.22, 0.09]	–0.70	.481
Performance	–1.23	0.07	[–1.36, –1.08]	–16.55	<.001
Speaker L1 × accentedness	–0.26	0.19	[–0.66, 0.12]	–1.36	.174
Speaker L1 × prestige	0.74	0.11	[0.52, 0.97]	6.70	<.001
Speaker L1 × performance	0.49	0.09	[0.30, 0.67]	5.14	<.001
Speaker L1 × prestige × performance	–1.01	0.13	[–1.28, –0.77]	–7.66	<.001
Control covariates					
Number of languages	5.70	10.01	[–14.52, 26.84]	0.52	.601
Accent familiarity	0.43	5.45	[–9.70, 10.30]	0.08	.938
Context familiarity	1.20	0.07	[1.06, 1.32]	17.44	<.001
Ethnic pride	2.98	6.12	[–7.78, 14.14]	0.49	.626

Random intercepts	**Variance**	***SD***
Listener	1.07	1.03
Speaker	0.07	0.27

Note. The reference level for speaker L1 is Brazilian Portuguese. The reference level for prestige is high. The reference level for performance is high.

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae042/7712341 by University of Chicago D'Angelo Law Library user on 13 July 2024

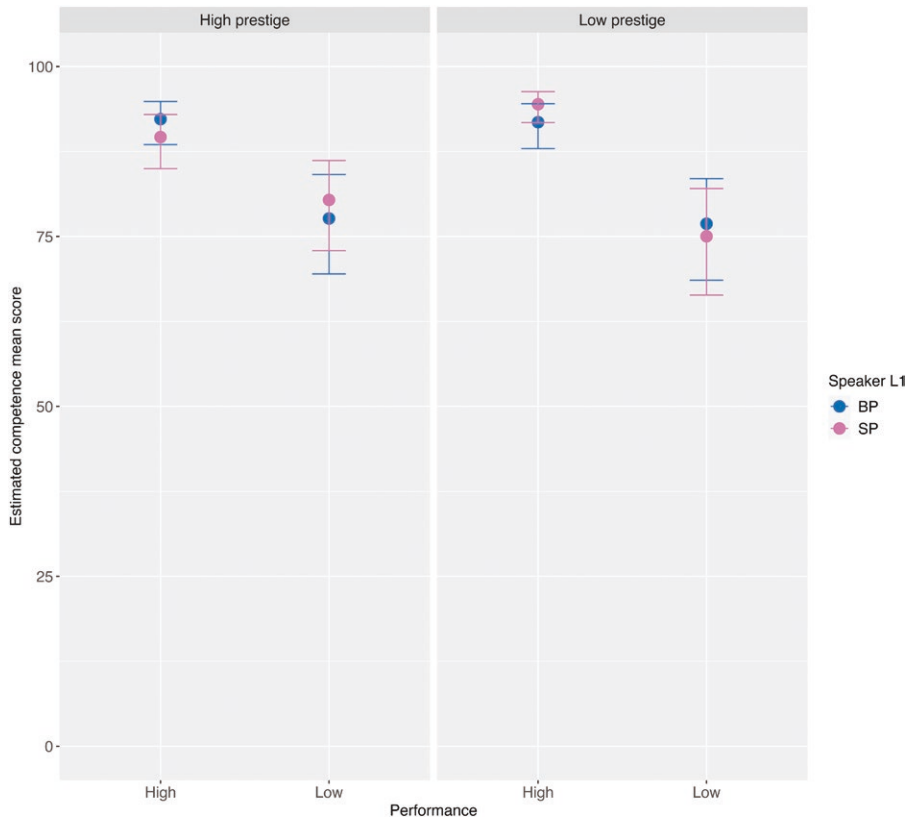

Figure 1: Estimated means for speaker competence as a function of speaker L1, job prestige, and performance level. Whiskers around estimated means enclose 95% CIs. BP = Brazilian Portuguese. SP = Spanish.

high, z = –9.91, $p < .001$, and in high-prestige jobs (over low-prestige jobs) when the performance level was low, z = 5.90, $p < .001$. Moreover, all speakers (regardless of their L1 and of the prestige level of the job) received statistically higher competence ratings when illustrating a high-level performance in comparison with low-level performance. However, job-relevant performance made a bigger difference to the competence ratings of L1 Spanish speakers, z = 26.17, $p < .001$, than to the ratings of L1 BP speakers, z = 17.00, $p < .001$, in low-prestige jobs, but it made more of a difference to the ratings of L1 BP speakers, z = 16.55, $p < .001$, than to the ratings of L1 Spanish speakers, z = 12.25, $p < .001$, in high-prestige jobs.

Treatment preference

As summarized in Table 2, for the ratings of listener treatment preference, there were significant main effects for speaker L1, accentedness, job prestige, and performance level. Overall, listeners gave higher preference ratings for L1 Spanish than L1 BP speakers, but also gave higher ratings to speakers who were rated as having a weaker accent in comparison to those whose accent was rated as stronger. Listeners generally assigned higher preference ratings to all speakers (regardless of their L1) in low-prestige jobs, while low-skill performances elicited lower treatment preference ratings overall. The effect for speaker L1 was qualified by an interaction with accentedness rating, where the difference in treatment preference ratings between L1 BP and Spanish speakers widened as the speaker was rated to have a stronger accent. The effect for speaker L1 was further

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae042/7712341 by University of Chicago D'Angelo Law Library user on 13 July 2024

Table 2: Summary of mixed-effects model for listener treatment preference

Fixed effects	Estimate	*SE*	95% CI	*z*	*p*
(Intercept)	0.30	0.59	[–0.89, 1.42]	0.52	.606
Speaker L1	1.20	0.28	[0.68, 1.74]	4.32	<.001
Accentedness	1.14	0.16	[0.78, 1.47]	6.96	<.001
Prestige	0.49	0.09	[0.32, 0.66]	5.70	<.001
Performance	–1.92	0.07	[–2.05, –1.78]	–27.27	<.001
Speaker L1 × accentedness	–0.87	0.19	[–1.26, –0.48]	–4.56	<.001
Speaker L1 × prestige	–0.21	0.12	[–0.45, 0.03]	–1.72	.086
Speaker L1 × performance	–0.22	0.10	[–0.38, –0.03]	–2.23	.026
Speaker L1 × prestige × performance	–0.66	0.14	[–0.95, –0.38]	–4.80	<.001
Control covariates					
Number of languages	–8.64	10.35	[–30.92, 13.02]	–0.83	.404
Accent familiarity	3.35	5.16	[–8.04, 13.64]	0.65	.516
Context familiarity	1.56	0.07	[1.41, 1.68]	22.38	<.001
Ethnic pride	2.29	5.81	[–8.60, 13.39]	0.40	.694

Random intercepts	Variance	*SD*
Listener	0.97	0.98
Speaker	0.07	0.26

Note. The reference level for speaker L1 is Brazilian Portuguese. The reference level for prestige is high. The reference level for performance is high.

qualified by an interaction with performance level, signifying that certain combinations of speaker L1 and performance level influenced listeners' evaluations differently. However, because some of these predictors are involved in a statistically significant three-way interaction among speaker L1, performance level, and job prestige (meaning that these three variables interacted with each other to influence treatment preference ratings; see full comparison details below), the interpretation of main effects and two-way interactions may be misleading. The only control covariate that emerged as significant was context familiarity, indicating that stronger context familiarity was associated with higher treatment preference ratings.

The three-way interaction among speaker L1, performance, and prestige (Figure 2) shows that in high-performance, high-prestige scenarios, L1 Spanish speakers received higher treatment preference ratings than BP speakers, $z = -2.62$, $p = .009$, while in low-performance, low-prestige scenarios, L1 BP speakers were the ones who were assigned higher ratings, $z = 2.00$, $p = .045$. L1 BP and Spanish speakers received comparable ratings in low-performance, high-prestige scenarios, $z = -1.80$, $p = .072$, and in high-performance, low-prestige scenarios, $z = -1.77$, $p = .077$. Beyond these direct comparisons between speaker groups, all speakers received higher ratings when in a low-prestige job than in a high-prestige job regardless of their L1 and performance level. The effect of prestige on treatment preference ratings of L1 BP, $z = -5.70$, $p < .001$, and L1 Spanish speakers, $z = -3.38$, $p < .001$, was similar when the performance level demonstrated by the speaker was high. However, when the speaker demonstrated low-level performance, the effect of job prestige was greater for L1 BP, $z = -18.98$, $p < .001$, than for L1 Spanish speakers, $z =$

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae042/7712341 by University of Chicago D'Angelo Law Library user on 13 July 2024

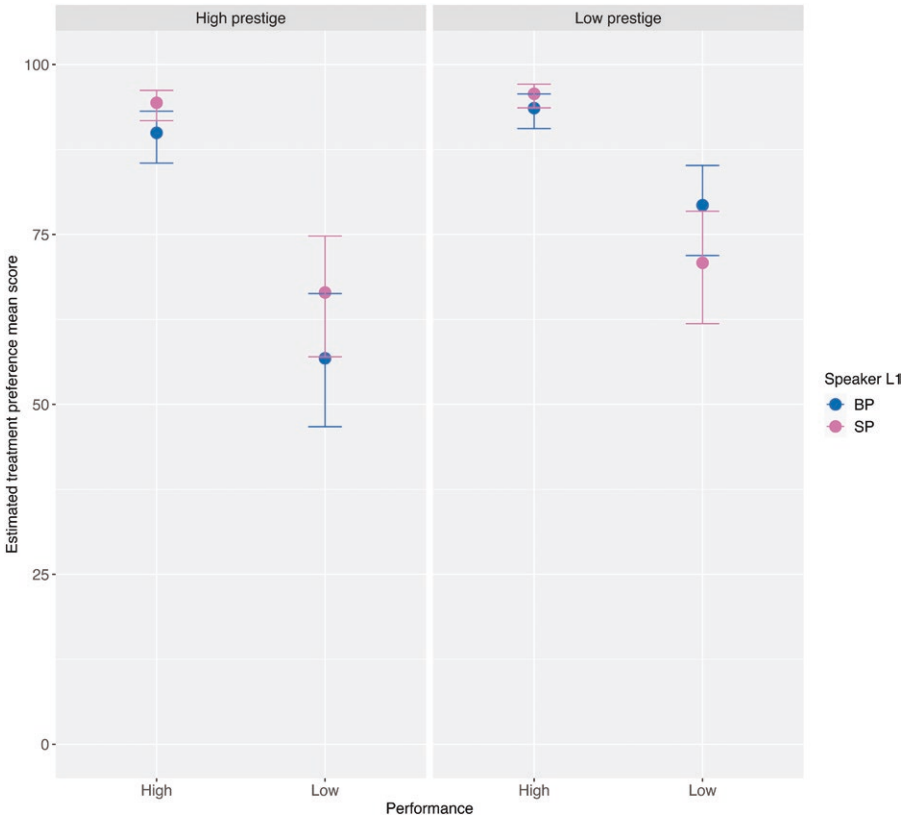

Figure 2: Estimated means for treatment preference as a function of speaker L1, job prestige, and performance level. Whiskers around estimated means enclose 95% CIs. BP = Brazilian Portuguese. SP = Spanish.

–3.92, $p < .001$. Similarly, all speakers received higher ratings when illustrating high-level performance than low-level performance regardless of their L1 and of the prestige of the job, but with notable differences in the effect of performance level. Overall, performance level made a bigger difference to the treatment preference ratings of the L1 Spanish speakers than to the ratings of the L1 BP speakers, with this difference being magnified in low-prestige scenarios, $z = 29.50$, $p < .001$ and $z = 18.21$, $p < .001$, respectively, in comparison to high-prestige scenarios, $z = 31.41$, $p < .001$ and $z = 27.27$, $p < .001$, respectively.

Comprehensibility

As summarized in Table 3, for the ratings of comprehensibility, there was only a significant main effect of accentedness, such that speakers rated as having weaker accents were judged as easier to understand than those rated as having stronger accents. The effect of speaker L1 was qualified by an interaction with accentedness, where the difference in comprehensibility ratings between BP and Spanish speakers widened as the speaker's accent was judged to be stronger. The effect of speaker L1 was further qualified by an interaction with prestige, signifying that the combination of these two variables affected comprehensibility ratings (see full comparison details below). Context familiarity approached significance at the five percent threshold, indicating that stronger context familiarity was associated with higher comprehensibility ratings.

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae042/7712341 by University of Chicago D'Angelo Law Library user on 13 July 2024

Table 3: Summary of mixed-effects model for speaker comprehensibility

Fixed effects	Estimate	SE	95% CI	z	p
(Intercept)	–0.15	0.97	[–2.13, 1.46]	0.16	.876
Speaker L1	–0.43	0.30	[–1.01, 0.13]	–1.43	.151
Accentedness	4.55	0.29	[4.05, 5.03]	15.82	<.001
Prestige	–0.17	0.16	[–0.46, 0.11]	–1.09	.274
Performance	0.04	0.16	[–0.35, 0.26]	–0.25	.805
Speaker L1 × accentedness	–0.65	0.32	[–1.13, –0.06]	–2.03	.042
Speaker L1 × prestige	0.65	0.17	[0.34, 0.98]	3.84	<.001
Speaker L1 × performance	–0.06	0.17	[–0.40, 0.27]	–0.37	.712
Speaker L1 × prestige × performance	0.02	0.22	[–0.44, 0.41]	0.09	.928
Control covariates					
Number of languages	18.24	17.57	[–14.58, 50.88]	1.04	.299
Accent familiarity	–1.40	8.83	[–18.23, 14.77]	–0.16	.874
Context familiarity	0.20	0.10	[–0.00, 0.38]	1.95	.051
Ethnic pride	3.16	9.89	[–15.08, 23.51]	0.32	.750

Random intercepts	Variance	SD
Listener	2.76	1.66
Speaker	0.03	0.16

Note. The reference level for speaker L1 is Brazilian Portuguese. The reference level for prestige is high. The reference level for performance is high.

The only significant interactions (Figure 3) that emerged in the comprehensibility model were between speaker L1 and accentedness and between speaker L1 and prestige. Notably, while L1 BP speakers received higher comprehensibility ratings when performing high-prestige jobs in comparison with low-prestige jobs, $z = 2.62$, $p = .009$, L1 Spanish speakers were judged more comprehensible when performing low-prestige jobs in comparison with high-prestige jobs, $z = –8.31$, $p < .001$. Further comparisons between the speaker groups in each prestige level revealed that in high-prestige scenarios, L1 BP speakers were perceived as significantly more comprehensible than L1 Spanish speakers, $z = 5.15$, $p < .001$, while both speaker groups received comparable comprehensibility ratings in low-prestige scenarios, $z = 1.56$, $p = .119$.

Discussion

We aimed to replicate and extend Teló *et al.* (2022), who showed that accent bias may be more nuanced than generally assumed by exploring the effects of job prestige and professional performance in the ratings of speaker competence, listener treatment preference, and speaker comprehensibility. We replicated the study in Brazil, a different linguistic setting that is marked by a smaller immigration influx, where the interactions between local- and foreign-born individuals are likely more limited in comparison to those commonly observed in Canada, and where the power dynamics between L1 and L2 speakers may be different from those experienced by interlocutors in Canada. Methodologically, we departed from the initial study in three key ways, one

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae042/7712341 by University of Chicago D'Angelo Law Library user on 13 July 2024

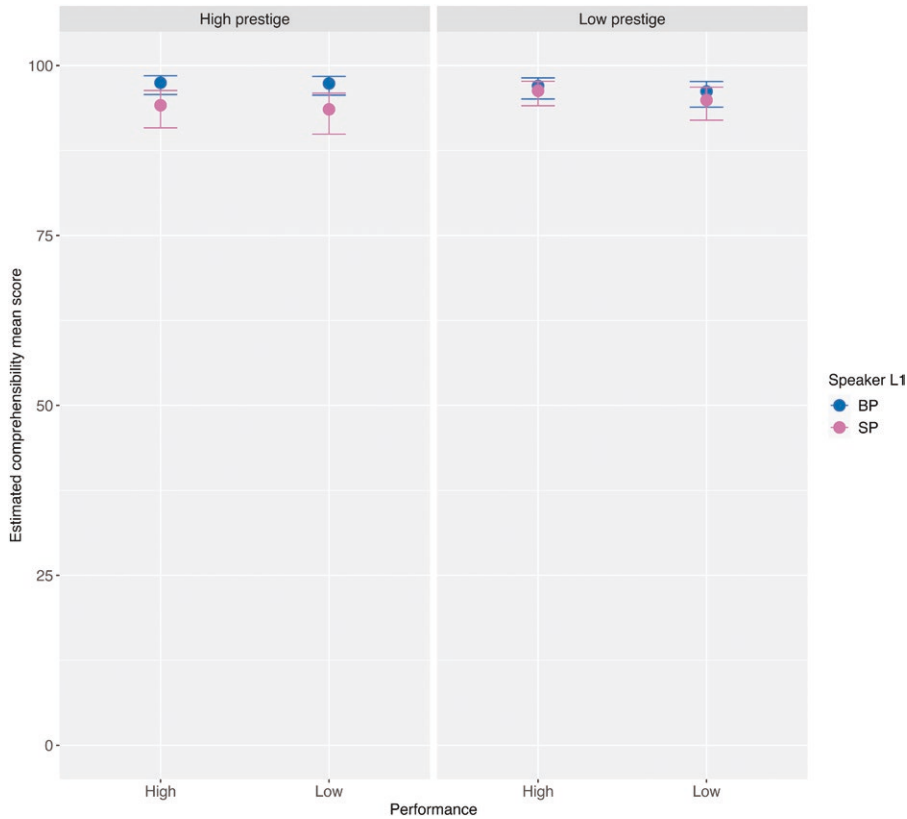

Figure 3: Estimated means for speaker comprehensibility as a function of speaker L1, job prestige, and performance level. Whiskers around estimated means enclose 95% CIs. BP = Brazilian Portuguese. SP = Spanish.

of which amounts to an extension of the initial study: First, we included only female speakers in the current replication. This decision aligns with recent studies on accent bias (e.g. Lindberg and Trofimovich 2023) and streamlined the procedures for data collection, allowing us to recruit fewer participants than Teló *et al.* (2022) while simultaneously collecting more ratings for each recording. Second, because it was important to ensure the coherence and relevance of the scripted content, our script adaptation process encompassed qualitative adjustments. Lastly, a notable addition to our analysis involved the inclusion of two new control covariates, one capturing listeners' familiarity with the context depicted in each script and another capturing listeners' ethnic pride.

Summary of findings

Collectively, our findings underscore the complicated nature of accent bias, revealing complexities that surpass the insights from the initial study. For instance, Teló *et al.* (2022) identified speaker L1 as a significant predictor across all models, signifying that listeners generally perceived L1 English speakers as more competent and comprehensible than L2 English speakers, and that listeners generally wished to be treated as presented in scripts recorded by L1 versus L2 speakers. Here, this significance was not replicated across the models examining competence and comprehensibility, suggesting that the weight tied to speaker L1 alone was diminished, resulting in listeners not uniformly perceiving L1 BP speakers as more competent and comprehensible

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae042/7712341 by University of Chicago D'Angelo Law Library user on 13 July 2024

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae042/7712341 by University of Chicago D'Angelo Law Library user on 13 July 2024

than L2 BP speakers. This difference across the studies is potentially related to the target languages, which in our study were Romance languages, likely contributing to mutual intelligibility. Furthermore, it may reflect the power differential between L1 BP and L1 Spanish speakers in Brazil, which is presumably different from the power imbalance observed between L1 English and L1 Tagalog speakers in Canada (Pennycook, 2010; Liddicoat, 2016). Additionally, in our treatment preference model, a divergence from the outcomes of the initial study surfaced: Listeners generally indicated a preference for treatment aligned with recordings from L2 BP speakers (see detailed discussion below). This unexpected twist implies that L1 background may interact with other (contextual, vocal) cues in unforeseen ways.

Beyond L1 per se, the listeners in the current replication exhibited a heightened effect of speakers' accentedness; in fact, this was the only variable that consistently emerged as significant across all models. Speakers who were rated as having weaker accents were judged as more competent and comprehensible compared to those who were rated as having stronger accents. This trend also extended to treatment preference ratings, where speakers with less prominent accents were favored. While this finding clearly implies a gradient (rather than categorical) influence of accent on evaluations (Ryan *et al.* 1977), it likely stems from the Spanish speakers' pronunciation posing little or virtually no difficulty of understanding to L1 BP-speaking listeners, given the degree of similarity between the phonological inventories of the two languages. It is possible that, because the L2 BP speakers were rated as comprehensible as the L1 BP speakers, listeners paid greater attention to (and were more critical of) speakers' accentedness, once it was the only linguistic dimension that set L1 and L2 speakers apart. Regardless, listeners favoring weaker accents over stronger accents (both in the initial study and here) is supported by previous research showing that customers prefer to interact with professionals with weaker accents (Tsunemoto *et al.* 2023). Additionally, it might suggest that listeners perceive more 'nativelike' accents (i.e. more phonetic-phonologically similar to the target language) as indicative of better communication proficiency (Derwing and Munro 2009). In turn, listeners may have upgraded these speakers assuming that they were professionals who possess valuable workplace characteristics, such as dynamism, sense of initiative, and leadership skills (Beaulieu *et al.* 2022). On a theoretical level, we may assume that the linguistic similarity between BP and Spanish yields high levels of speaker comprehensibility, which, according to the processing fluency account of accent discrimination (Dragojevic and Giles 2016), should result in lower levels of accent discrimination overall in the present study when compared to the initial study. While our results may, indeed, demonstrate this to a certain extent, it is noteworthy that even though the listeners perceived the L1 and the L2 speakers comparably comprehensible, they showed clear accent-based biases. This provides support for categorization and stereotyping accounts of accent discrimination (e.g., Ryan 1983).

Lastly, reflecting the built-in manipulation of job-relevant performance level, listeners rated all speakers (irrespective of their L1 and of the prestige level of the job) as more competent in high- versus low-level performance scenarios. Similarly, regardless of speakers' L1 and the prestige of the job, listeners expressed a stronger preference to be treated in accordance with the high-performance scenarios rather than the low-performance counterparts. These findings, which were generally in line with those of the initial study, were obtained after we statistically controlled four variables capturing individual differences in the number of languages known by the listeners and their accent familiarity (as in Teló *et al.* 2022), and their context familiarity and ethnic pride. In what follows, we discuss the findings that are meaningful when the full dataset is considered (i.e. based on the three-way interactions, whenever they were statistically significant).

Competence and treatment preference

Beyond the patterns described above, the relative effects of performance level and job prestige on listeners' ratings reveal important distinctions between L1 and L2 speakers, including in the types of jobs and performances that listeners associate with L1 and L2 speakers, which constitutes linguistic bias. For example, regarding competence, when L1 Spanish speakers demonstrated

high-level performance, they elicited higher competence ratings in low-prestige jobs than in high-prestige jobs. This might have tapped into listeners' assumptions about L2 speakers (e.g. Russo *et al.* 2017), such that they were perceived as more competent in low-stakes, low-prestige jobs even when other L2 speakers were depicted as skillful professionals in high-status jobs. Similarly, hearing L2 speakers in low-status jobs potentially tapped into listeners' expectations about or experience with immigrants doing low-status, low-paying jobs well—as reported in previous research in Brazil, for example, Tedesco and Grzybrovski 2012—leading to higher competence ratings. These patterns aside, when L2 speakers' performance level was low, they were rated more competent in high-prestige scenarios compared to low-prestige scenarios, perhaps because individuals require greater competence (e.g. success in education) to attain the high-prestige jobs included in the study (doctor, lawyer, professor). It is noteworthy that none of this was true for L1 BP speakers, whose competence ratings remained comparable across the different scenarios, with only a high- versus low-performance distinction.

In a greater degree of nuance than that discussed in Teló *et al.* (2022), our analyses further show that job-relevant performance affected listeners' ratings differently depending on the speaker's L1 and the prestige of the job. In summary, when speakers took on the role of a professional in a low-prestige job, performance made a bigger difference to the competence and treatment preference ratings of L2 speakers than to the ratings of L1 speakers. Conversely, when job prestige was high, the impact of performance on perceived competence (but not treatment preference) was more pronounced for L1 speakers. Broadly, this provides additional evidence that listeners may have different expectations for L1 and L2 speaker (Lindemann and Subtirelu 2013), including in the workplace (Russo *et al.*, 2017), creating a bias in how performance is perceived, evaluated, and, perhaps, processed (Ridgway 1991; Foschi 2013), especially in jobs with varying prestige levels (Kalin and Rayko 1978; Singer and Eder 1989). In the case of competence specifically, this finding indicates that listeners may hold different competence standards, applying stricter or more lenient performance standards to individuals based on factors such as their ethnicity, place of origin, and socioeconomic class, as well as based on the social interaction context (e.g. Biernat and Kobrynowicz 1997; Foschi 2000).

Regarding listener treatment preference specifically, these ratings showed a more straightforward effect of speaker L1, with L2 BP speakers generally receiving higher preference ratings than L1 BP speakers, which is strikingly different from the results of the initial study. This finding may imply that listeners valued certain communication styles and/or pronunciation features associated with the L2 speakers, which is underscored by communication- and speech-related comments provided by listeners to elucidate their evaluation of scenarios involving low-level performance presented by Spanish speakers. Comments such as 'She lacks knowledge about her field of work, but she's polite and friendly' and 'This person uses a very graceful way of speaking' indicate that listeners may have drawn upon various linguistic elements, including pronunciation but extending beyond mere accent, to formulate their treatment preference judgments. This phenomenon is akin to how listeners often associate certain speech features with personal traits, leading to diverse behavioral responses toward those speakers (e.g., Fasoli *et al.* 2017; Pearsell and Pape 2023). Other comments suggest that listeners may have been positively influenced by the proficiency level exhibited by the L2 BP speakers, thus rewarding them in comparison to L1 BP speakers, whose 'nativelike' linguistic performance was already expected and thus not surprising (Teló *et al.* 2024). This is demonstrated, for instance, in a comment provided by a listener upon hearing a recording of an L2 BP speaker: 'Polite vocabulary. The message has a logical continuity that proves good reasoning and fluency. It gives the impression that she thinks in Portuguese, almost like a native speaker. The use of an expression in the diminutive shows that she is used to the popular lexicon of Brazil, which gives an even greater tone of intimacy with the Portuguese language'. Regardless of the explanation, prior research has obtained similar findings, where speakers who are rated lower on both pronunciation proficiency measures (accentedness, comprehensibility, fluency) and status and solidarity traits (e.g. honesty, education) also receive higher professional-related preference ratings from listeners (Trofimovich and Turuševa 2020).

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae042/7712341 by University of Chicago D'Angelo Law Library user on 13 July 2024

A noteworthy exception to this pattern was when both prestige and performance level were low. In these scenarios, listeners shifted their preference toward treatment by L1 speakers, which was likely driven by the greater impact that job prestige had on the treatment preference ratings assigned to L1 over L2 speakers when low-level performance was demonstrated. Again, this is opposite to what is reported in Teló *et al.* (2022), where the sole circumstance favoring treatment by L2 speakers was low-prestige, low-performance scenarios. Nonetheless, our listeners' behavior seems to be in line with a robust body of evidence showing not only people's strong preference for highly skilled immigrants over low-skilled immigrants but also opposition to low-skilled immigration in general (e.g. Hainmuller and Hiscox 2010).

Lastly, additional findings further expand our understanding of the effect of job prestige on listeners' treatment preference. The ratings for treatment preference indicated a general preference for interaction styles associated with low-prestige jobs, irrespective of speakers' L1 or performance. While identifying the precise mechanisms underlying this interaction is out of the scope of the present research, it is possible that listeners simply prefer the lower degree of formality and the higher levels of approachability likely conveyed by the speakers in those roles (e.g. Fast and Funder 2008). Similarly, listeners may feel a reduced social distance with speakers who employ interaction and/or speech styles prevalent in less formal, low-prestige jobs, fostering a sense of ease and openness in communication (Heylighen and Dewaele 2002; Fast and Funder 2008). More broadly, this could also be related to cultural differences in how people in Brazil and Canada consider job status/prestige and react to L2 speech.

Comprehensibility

The results from the comprehensibility model offer valuable insights into how different factors influence listeners' perception of ease of understanding. The most notable and consistent effect observed in the comprehensibility ratings is the influence of accentedness, where speakers who received lower accentedness evaluations were generally perceived as more comprehensible than those who received higher accentedness ratings. This aligns with both the initial study's findings and a well-established body of research (Derwing and Munro 2015). The effect of speaker L1 on comprehensibility, although not statistically significant, was qualified by an interaction with accentedness, signaling that the difference in comprehensibility ratings between BP and Spanish speakers widened as the speaker's accent became stronger.

The interplay between speaker L1 and comprehensibility takes on greater nuance in conjunction with prestige. The interaction between speaker L1 and prestige implies that the relationship between comprehensibility and job prestige is different for L1 BP and Spanish speakers, where basically the same L1 Spanish speakers were judged more comprehensible in low- versus high-prestige jobs, while L1 BP speakers were rated more comprehensible in high-prestige scenarios than in low-prestige scenarios. Hearing L1 Spanish speakers in low-prestige occupations and L1 BP speakers in high-prestige roles likely tapped into expectations held by the listeners, such that L2-accented speakers are more fitting for low-prestige occupations (e.g., Kalin and Rayko 1978). Importantly, for L2 speakers, comprehensibility ratings generally followed competence ratings, with L1 Spanish speakers being rated as more comprehensible and more competent in low- versus high-prestige scenarios. Regardless of the explanation, listeners appeared to conflate pronunciation with professional competence for L2 speakers but did so to a lesser degree when assessing L1 speakers (Hosoda *et al.* 2007; Dragojevic and Giles 2016). Finally, differently from the initial study, speakers' demonstrated performance level did not seem to interact with the comprehensibility ratings in a meaningful way, suggesting a less complicated picture for comprehensibility.

Listener variables as predictors of speaker ratings

Considering prior research indicating the potential influence of various listener-related factors on L2 speech assessment, we expanded the individual differences analysis reported in the initial study. We modeled four control covariates: the number of languages known by the listeners, their

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae042/7712341 by University of Chicago D'Angelo Law Library user on 13 July 2024

familiarity with L2-accented speech, their familiarity with the context in each script, and their ethnic pride. Despite a noteworthy difference between the listener samples in Teló *et al.* (2022) and here, particularly with all listeners in the current replication being L1 BP speakers as opposed to 73% of L1 English-speaking listeners in the initial study, the number of languages known by the listeners and their familiarity with L2-accented speech did not interact with listeners' ratings, as in Teló *et al.* (2022). Context familiarity featured as a significant predictor of listeners' ratings of competence, treatment preference, and comprehensibility ($p = .051$), where greater context familiarity was consistently related to higher ratings (Gass and Varonis 1984). This is an important finding of the present study, and we encourage more research to understand the mechanisms of this interaction, which could be, for example, rooted in empathy (familiarity with the context might foster a sense of empathy and connection between the listener and the speaker), expectation alignment (listeners who are familiar with the context might have aligned expectations with the speaker's communication style, leading to a more positive reception of the speaker and their message), or cognitive ease (familiarity with the context may reduce cognitive load for listeners). Nonetheless, it is important to acknowledge that the question aiming to elicit listeners' context familiarity could be interpreted in slightly different ways (e.g. the listener could or not consider speaker-related information, such as their gender and L1, in determining how familiar they are with that specific context). Despite the robust effect that the variable demonstrated in our models, some caution is necessary when generalizing this finding. Finally, ethnic pride did not seem to be related to the ratings meaningfully.

Limitations, implications, and future work

As a close replication, this study shares limitations with the original study, such as the fact that a detailed account of accent bias in any context would require systematic comparisons of representative samples of diverse listener groups. In addition, listeners did not have access to speakers' visual cues. While our study extends the original investigation to novel ethnolinguistic backgrounds, our findings are specific to the linguistic pairing of Spanish and Brazilian Portuguese, which exhibit a high degree of similarity. As discussed above, our results suggest that processing L1- and L2-accented speech was similarly easy for the listeners. While this provides novel insights into the phenomenon of accent bias by reporting on evaluations provided to L2-accented speakers whose L2 pronunciation was generally not linked to lower comprehensibility, the cultural closeness between L1 BP and L1 Spanish speakers should also be considered, which is likely not the same for L1 English speakers from Canada and L1 Tagalog speakers from the Philippines. Relatedly, while there may be a power imbalance between L1 BP and L1 Spanish speakers in Brazil, the global context does not reinforce this imbalance to the same degree as it does to the power differential between L1 English and L1 Tagalog speakers in Canada. We encourage researchers to investigate linguistic and cultural pairings with varying degrees of proximity and power dynamics in future research. Additionally, it is essential to acknowledge that our results may not reflect the evaluations of other immigrant groups in Brazil, such as Haitians, who face a greater likelihood of eliciting negative perceptions in comparison to other groups like Latinxs or Europeans (Mundin and Santos 2022). Similarly, given the sociopolitical and sociolinguistic landscapes of Santa Catarina, listeners from other states may hold different views toward the speakers. We also encourage researchers to adapt the methodological design for qualitative studies given that ratings provide only general patterns, which we might not be able to interpret accurately. To inspire future research, we provide a summary of the comments provided by the listeners as they evaluated the audio recordings (see Supplementary Data and Information). These data could help, among other things, to identify listener variables that deserve more attention in accent bias research. Furthermore, because voice features that were not controlled for may have influenced listeners' evaluations, especially of the professional and experiential constructs, further research is needed to disentangle the relative contributions of accent- and other pronunciation-related features (e.g. voice quality, speaking style) to listener evaluations. Lastly, our study used scripts that featured qualitative changes (see Method section) compared to the scripts used in Teló *et al.* (2022), potentially contributing to the obtained results.

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae042/7712341 by University of Chicago D'Angelo Law Library user on 13 July 2024

Despite their overall promising outlook, our results point to the presence of accent bias in listeners' assessments. This emphasizes the need for dismantling unjustified associations between a person's accent and their qualifications, as well as long-standing notions that certain forms of self-expression and communicative practices are more or less compatible with an organization's dominant culture (Park 2020). Achieving these societal goals requires critical education. The finding that listeners' context familiarity was positively related to their ratings also holds implications: If listener familiarity with speakers in certain roles matters, listener interventions aimed at reducing accent bias in professional settings should expose listeners to more examples of L2 speakers in high-performance, high-prestige contexts to set the expectation that they will encounter such speakers in these situations in the future. Similarly, more representation of L2 speakers in high-performance, high-prestige contexts in popular media should be advocated for (Eagly and Koenig 2021). In summary, addressing accent bias and promoting fair evaluations in professional communication environments requires a collective effort that encompasses education and the reevaluation of preconceived notions about language and language users.

Supplementary data

Supplementary material is available at *Applied Linguistics* online.

Notes on Contributors

Cesar Teló holds a master's degree in Applied Linguistics from the University of Calgary, Alberta, Canada. He is broadly interested in second language speech, including how learners acquire and process individual sounds, how individual differences affect the learning process, how listeners react to second language speech, how speaking with a second language accent impacts speakers' lives, and how pronunciation can be taught and learned more critically and effectively. *Address for correspondence*: School of Languages, Linguistics, Literatures and Cultures, University of Calgary, 2500 University Dr. NW, Calgary, AB T2N 1N4, Canada. <cesar.telo@ucalgary.ca>

Rosane Silveira is a faculty member of the Department of Foreign Language and Literature and the English Graduate Program at Universidade Federal de Santa Catarina, Brazil. CNPq research fellow. Her research interests include the development of second language speech, with a focus on phonetics, phonology and pronunciation teaching. <rosane@cce.ufsc.br>

Ana Flávia Boeing Marcelino is a visiting research associate at the Department of Romance Languages and Literatures at the University of Chicago, USA. She holds a master's degree in English Linguistic and Literary Studies from Universidade Federal de Santa Catarina, Brazil. Her research interests include the teaching and learning of Brazilian Portuguese, Task-based Language Teaching, material design and Brazilian culture, and Domain Analyses for Global Business Communication. Her most recent publication addresses anticolonialist practices in the Portuguese as a Host Language classroom. <anaflavia@uchicago.edu>

Mary G. O'Brien is an Adjunct Professor of German at the University of Calgary and Professor of Linguistics at Simon Fraser University, where she also serves as Dean and Vice-Provost of Graduate Studies. Her research focuses on the perception and production of L2 speech. In her recent, highly collaborative work, she and her colleagues have examined the role of a range of social factors and methodological issues in speech rating research. She is especially interested in the real-world implications of speaking with an accent. <mary_obrien@sfu.ca>

Notes

[1] Although some previous studies have demonstrated that reactions to the speech of male and female speakers may differ, the study replicated here did not demonstrate differences.
[2] As pointed out by an anonymous reviewer, since we controlled for the type of L2 accent among speakers, it would be relevant to model listeners' familiarity with that particular L2 accent. We encourage researchers to collect this information in future studies.

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae042/7712341 by University of Chicago D'Angelo Law Library user on 13 July 2024

References

Bagno, M. 1999. *Preconceito linguístico: O que é, como se faz*. Edições Loyola.

Baquiran, C. L. C. and E. Nicoladis. 2020. 'A doctor's foreign accent affects perceptions of competence,' *Health Communication* 35/6: 726–30. doi: 10.1080/10410236.2019.1584779

Bates, D., *et al*. 2015. 'Fitting linear mixed-effects models using lme4,' *Journal of Statistical Software* 67/1: 1–48. doi: 10.18637/jss.v067.i01

Baum, C. F. 2008. 'Stata tip 63: Modeling proportions,' *The Stata Journal* 8/2: 299–303. doi: 10.1177/1536867X0800800212

Beaulieu, S., *et al*. 2022. 'Professional identities of French Lx economic immigrants: Perceptions from a local French-speaking community,' *Languages* 7/2: 140–15. doi: 10.3390/languages7020140

Biernat, M. and D. Kobrynowicz. 1997. 'Gender- and race-based standards of competence: Lower minimum standards but higher ability standards for devalued group,' *Journal of Personality and Social Psychology* 72/3: 544–57. doi: 10.1037//0022-3514.72.3.544

Bueno, A. M. 2018. 'Norma linguística, preconceito e imigração no Brasil,' *Linguagem & Ensino* 21/1: 427–47. https://periodicos-old.ufpel.edu.br/ojs2/index.php/rle/article/view/15160.

Cavalcanti, L., T. Oliveira, and B.G. Silva. 2022. *Relatório anual OBMigra 2022*. Ministério da Justiça e Segurança Pública, Observatório das Migrações Internacionais. https://portaldeimigracao.mj.gov.br/pt/dados/relatorios-a.

Derwing, T.M. and M.J. Munro. 2015. *Pronunciation Fundamentals: Evidence-Based Perspectives for L2 Teaching and Research*. John Benjamins. doi:10.1075/lllt.42

Derwing, T. M. and M. J. Munro. 2009. 'Putting accent in its place: Rethinking obstacles to communication,' *Language Teaching* 42/4: 476–90. doi: 10.1017/S026144480800551X

Dragojevic, M. and H. Giles. 2016. 'I don't like you because you're hard to understand: The role of processing fluency in the language attitudes process,' *Human Communication Research* 42/3: 396–420. doi: 10.1111/hcre.12079

Eagly, A. H. and A. M. Koenig. 2021. 'The vicious cycle linking stereotypes and social soles,' *Current Directions in Psychological Science* 30/4: 343–50. doi: 10.1177/09637214211013775

Fasoli, F., *et al*. 2017. 'Gay- and lesbian-sounding auditory cues elicit stereotyping and discrimination,' *Archives of Sexual Behavior* 46/5: 1261–77. doi: 10.1007/s10508-017-0962-0

Fast, L. A. and D. C. Funder. 2008. 'Personality as manifest in word use: Correlations with self-report, acquaintance report, and behavior,' *Journal of Personality and Social Psychology* 94/2: 334–46. doi: 10.1037/0022-3514.94.2.334

Foschi, M. 2000. 'Double standards for competence: Theory and research,' *Annual Review of Sociology* 26: 21–42. https://www.jstor.org/stable/223435

Foschi, M. 2013. 'The professional credentials of immigrants: A status-and-expectations approach' in Thye, S.R. and E.J. Lawler (eds): *Advances in Group Processes*. Emerald, pp. 131–55. doi:10.1108/S0882-6145(2013)0000030009

Gass, S. and E. M. Varonis. 1984. 'The effect of familiarity on the comprehensibility of nonnative speech,' *Language Learning* 34/1: 65–87. doi: 10.1111/j.1467-1770.1984.tb00996.x

Hainmuller, J. and M. J. Hiscox. 2010. 'Attitudes toward highly skilled and low-skilled immigration: Evidence from a survey experiment,' *American Political Science Review* 104/1: 61–84. doi: 10.1017/S0003055409990372

Hamel, R.E. 2013. 'Language policy and ideology in Latin America' in Bayley, R., R. Cameron, and C. Lucas (eds.): *The Oxford handbook of sociolinguistics*. Oxford University Press, pp. 609–28. doi: 10.1093/oxfordhb/9780199744084.013.0030

Heylighen, F. and J. -M. Dewaele. 2002. 'Variation in contextuality of language: An empirical measure,' *Foundations of Science* 7/3: 293–340. doi: 10.1023/A:1019661126744

Hosoda, M., E. F. Stone-Romero, and J. N. Walter. 2007. 'Listeners' cognitive and affective reactions to English speakers with standard American English and Asian accents,' *Perceptual and Motor Skill* 104/1: 307–26. doi: 10.2466/pms.104.1.307-326

Kalin, R. and D. S. Rayko. 1978. 'Discrimination in evaluative judgments against foreign-accented job candidates,' *Psychological Reports* 43/3: 1203–9. doi: 10.2466/pr0.1978.43.3f.1203

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae042/7712341 by University of Chicago D'Angelo Law Library user on 13 July 2024

Kuznetsova, A., P. B. Brockhoff, and R. H. B. Christensen. 2017. 'lmerTest package: Tests in linear mixed effects models,' *Journal of Statistical Software* 82/13: 1–26. doi: 10.18637/jss.v082.i13

Lenth, R. 2023. *emmeans: Estimated marginal means, aka least-squares means*. R package version 1.8.7. https://CRAN.R-project.org/package=emmeans

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

Liddicoat, A. J. 2016. 'Native and non-native speaker identities in interaction: Trajectories of power,' *Applied Linguistics Review* 7/4: 409–29. doi: 10.1515/applirev-2016-0018

Lindberg, R. and P. Trofimovich. 2023. 'When accent does not match expectations: A dynamic perspective of L2 speaker evaluations in a French interview context,' *Canadian Journal of Applied Linguistics* 26/2: 1–31. doi: 10.37213/cjal.2023.33052

Lindemann, S. 2003. 'Koreans, Chinese or Indians? Attitudes and ideologies about nonnative English speakers in the United States,' *Journal of Sociolinguistics* 7/2: 348–64. doi: 10.1111/1467-9481.00228

Lindemann, S. and N. Subtirelu. 2013. 'Reliably biased: The role of listener expectation in the perception of second language speech,' *Language Learning* 63/3: 567–94. doi: 10.1111/lang.12014

Lippi-Green, R. 2012. *English with an accent: Language, ideology, and discrimination in the United States*. Routledge. doi: 10.4324/9780203348802

Marsden, E., *et al*. 2018. 'Replication in second language research: Narrative and systematic reviews and recommendations for the field,' *Language Learning* 68/2: 321–91. doi: 10.1111/lang.12286

Mendes de Oliveira, M. 2015. "Assertiveness, compliance, and politeness: Pragmatic and sociocultural aspects of 'Brazilian English' and 'American English", *International Journal of Society, Culture & Language* 3/1: 76–90. https://www.ijscl.com/article_11538.html.

Mundin, P. S. and C. S. R. Santos. 2022. 'Preconceito racial ou competição econômica? A opinião pública sobre a vinda de estrangeiros para o Brasil,' *Colombia Internacional* 110: 123–44. doi: 10.7440/colombiaint110.2022.05

Park, J. S. -Y. 2020. 'Translating culture in the global workplace: Language, communication, and diversity management,' *Applied Linguistics* 41/1: 109–28. doi: 10.1093/applin/amz019

Pearsell, S. and D. Pape. 2023. 'The effects of different voice qualities on the perceived personality of a speaker,' *Frontiers in Communication* 7: 1–20. doi: 10.3389/fcomm.2022.909427

Pennycook, A. 2010. *Language as a local practice*. Routledge. doi:10.4324/9780203846223

R Core Team. 2023. *R: A language and environment for statistical computing (Version 4.3.0)*. R Foundation for Statistical Computing. https://www.R-project.org/

Revelle, W. 2023. *psych: Procedures for psychological, psychometric, and personality research*. R package version 2.3.6. https://CRAN.R-project.org/package=psych

Ridgway, C. 1991. 'The social construction of status value: Gender and other nominal characteristic,' *Social Forces* 70/2: 367–86. doi: 10.1093/sf/70.2.367

Rinke, S. 2014. 'The reconstruction of national identity: German minorities in Latin America during the First World War' in Foote N. and M. Goebel (eds.): *Immigration and national identities in Latin America*. University Press of Florida, pp. 160–81.

Rodrigues, C.C. 2014. 'Língua no Brasil: Variação e multilinguismo' in Rodrigues, C.C, T.R. de Luca, and V. Guimarães (eds.): *Identidades brasileiras: Composições e recomposições*. Cultura Acadêmica, pp. 117–44. http://hdl.handle.net/11449/123650.

Russo, M., G. Islam, and B. Koyuncu. 2017. 'Non-native accents and stigma: How self-fulfilling prophesies can affect career outcomes,' *Human Resource Management Review* 27/3: 507–20. doi: 10.1016/j.hrmr.2016.12.001

Ryan, E. B. 1983. 'Social psychological mechanisms underlying native speaker evaluations of non-native speech,' *Studies in Second Language Acquisition* 5/2: 148–59. doi: 10.1017/S0272263100004824

Ryan, E. B., M. A. Carranza, and R. W. Moffie. 1977. 'Reactions toward varying degrees of accentedness in the speech of Spanish-English bilinguals,' *Language and Speech* 20/3: 267–73. doi: 10.1177/002383097702000308

Seyferth, G. 2011. 'A dimensão cultural da imigração,' *Revista Brasileira de Ciências Sociais* 26/77: 47–62. doi: 10.1590/S0102-69092011000300007

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae042/7712341 by University of Chicago D'Angelo Law Library user on 13 July 2024

Singer, A. 2021. 'A reativação da direita no Brasil,' *Opinião Pública* 27/3: 705–29. doi: 10.1590/1807-01912021273705

Singer, M. and G. S. Eder. 1989. 'Effects of ethnicity, accent, and job status on selection decisions,' *International Journal of Psychology* 24/1–5: 13–34. doi: 10.1080/00207594.1989.10600029

Spence, J. L., *et al.* 2024. 'Is your accent right for the job? A meta-analysis on accent bias in hiring decisions,' *Personality and Social Psychology Bulletin* 50/3: 371–86. doi: 10.1177/01461672221130595

Taylor Reid, K., P. Trofimovich, and M. G. O'Brien. 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: 419–42. doi: 10.1017/S0272263118000244

Tedesco, J. C. and D. Grzybovski. 2012. 'Senegaleses no norte do Rio Grande do Sul: Integração cultural, trabalho e dinâmica migratória internacional,' *Revista Espaço Pedagógico* 18/2: 336–55. https://seer.upf.br/index.php/rep/article/view/2433.

Teló, C., P. Trofimovich, and M. G. O'Brien. 2022. 'Disentangling professional competence and foreign accent,' *Journal of Second Language Pronunciation* 8/3: 413–43. doi: 10.1075/jslp.22013.tel

Teló, C., *et al.* 2024. 'Beyond the resume: HR students' evaluations of interview performances by first and second language speakers. Manuscript submitted for publication.

Thomas, J. 1983. 'Cross-cultural pragmatic failure,' *Applied Linguistics* 4/2: 91–112. doi: 10.1093/applin/4.2.91

Trofimovich, P. and L. Turuševa. 2020. 'Language attitudes and ethnic identity: Examining listener perceptions of Latvian-Russian bilingual speakers,' *Journal of Language, Identity & Education* 19/1: 9–24. doi: 10.1080/15348458.2019.1696682

Tsunemoto, A., *et al.* 2023. 'Listener perceptions of customer service agents' performance: Do comprehensibility, accentedness, and fluency matter?,' *Journal of Second Language Pronunciation* 9/2: 234–62. doi: 10.1075/jslp.21027.tsu

Downloaded from https://academic.oup.com/applij/advance-article/doi/10.1093/applin/amae042/7712341 by University of Chicago D'Angelo Law Library user on 13 July 2024