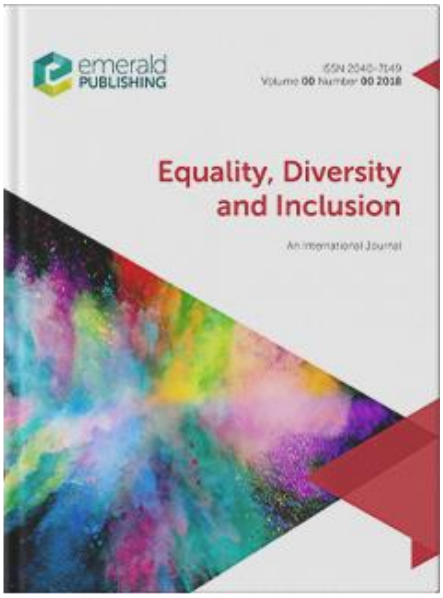

Beyond the resume: HR students' evaluations of interview performances by first and second language speakers

Cesar Teló, Pavel Trofimovich, Mary G. O'Brien, Thao-Nguyen Nina Le, and Anamaria Bodea

Teló, C., Trofimovich, P., O'Brien, M. G., Le, T.-N. N., Bodea, A. (2024). Beyond the resume: HR students' evaluations of interview performances by first and second language speakers. *Equality, Diversity and Inclusion: An International Journal*. Published online 27 August 2024. https://doi.org/10.1108/EDI-10-2023-0335

Beyond the resume: HR students' evaluations of interview performances by first and second language speakers

Abstract

Purpose

High-stakes decision-makers, including human resource (HR) professionals, often exhibit accent biases against second language speakers in professional evaluations. We extend this work by investigating how HR students evaluate simulated job interview performances in English by first and second language speakers of English.

Design/methodology/approach

Eighty HR students from Calgary and Montreal evaluated the employability of first language (L1) Arabic, English, and Tagalog candidates applying for two positions (nurse, teacher) at four points in the interview (after reading the applicant's resume, hearing their self-introduction, and listening to each of two responses to interview questions). Candidates' responses additionally varied in the extent to which they meaningfully answered the interview questions.

Findings

Students from both cities provided similar evaluations, employability ratings were similar for both advertised positions, and high-quality responses elicited consistently high ratings while evaluations for low-quality responses declined over time. All speakers were evaluated similarly based on their resumes and self-introductions, regardless of their language background. However, evaluations diverged for interview responses, where L1 Arabic and Tagalog speakers were considered more employable than L1 English speakers. Importantly, students' preference for L1 Arabic and Tagalog candidates over L1 English candidates was magnified when those candidates provided low-quality interview responses.

Originality

Results suggest that even in the absence of dedicated equity, diversity, and inclusion (EDI) training focusing on language and accent bias, HR students may be aware of second language speakers' potential disadvantages in the workplace, rewarding them in the current evaluations. Findings also highlight the potential influence of contextual factors on HR students' decision-making.

Keywords

Accent bias, employability, hiring discrimination, microaggressions, decision making, HR students, EDI agenda, interview sequence, comprehensibility.

Introduction

From an EDI (equity, diversity, and inclusion) perspective, human resource (HR) professionals, who are tasked with making high-stakes decisions, may show biases. For the purposes of this work, we focus on accent-based biases against second language (L2) speakers whose L2 accent is influenced by their previously known languages (henceforth referred to as accented L2 speakers).[1] Such biases, which can be described as microaggressions (Trofimovich *et al*., 2024) or microinequities, can have significant real-world repercussions for accented L2 speakers, including systematic downgrading in professional evaluations (Spence *et al*., 2024). Despite a rapidly changing immigration landscape and increased awareness of discrimination, university-level HR training programs in Canada—the context of the present study—do not typically address linguistic diversity or accent bias (i.e., discrimination on the basis of an L2 accent) through their curricula (O'Brien *et al*., 2024). In this study, we therefore investigated accent bias among university-level HR students in two Canadian cities. We examined how HR

students (henceforth referred to as students) evaluate mock job interview performances of first language (L1) Arabic, English, and Tagalog speakers while exploring the roles of job type (teacher vs. nurse), response quality (high vs. low), and timing of assessment (before an applicant is heard vs. at various points in the interview) in their hiring decisions. Arabic and Tagalog were chosen as the target L2 accents because of the demographic relevance of these L1s in the context of this investigation.

Background Literature

In contemporary societies, discrimination has typically shifted away from overt, blatant acts toward more subtle forms known as microaggressions, which encompass brief, everyday slights and insults conveying negative messages (Sue *et al.*, 2007). Within the realm of microaggressions, language- and accent-related instances of microinsults and microinvalidations are particularly prevalent (Trofimovich *et al.*, 2024), including in the workplace (e.g., Näre and Cleland Silva, 2021). Characterized by their subtlety and insensitivity, microinsults include, for example, the mocking of accented L2 speakers, with their accents serving as the basis for attributing them lower competence. Microinvalidations are marked by situations where individuals are excluded or their experiences minimized, such as when students receive praise for their exceptional language skills, concealing an underlying message of surprise that someone from an ethnic minority could excel in the majority language.

A substantial body of research has scrutinized how different groups are discriminated against, with clear implications for EDI research and practices (e.g., Comer *et al.*, 2023). With respect to accented L2 speakers in particular, they tend to receive less favorable social, experiential, and professional evaluations (e.g., perceived as less trustworthy, credible, effective, and competent) than L1 speakers (e.g., Baquiran and Nicoladis, 2020; Lev-Ari and Keysar, 2010;

Ramjattan, 2021; Teló *et al*., 2022). These negative perceptions hold particular significance within decision-making domains, where professionals serve as gatekeepers to employment opportunities. Unfavorable evaluations frequently translate into diminished prospects for hiring, promoting, and retaining accented L2 speakers (e.g., Ertorer *et al*., 2020; Näre and Cleland Silva, 2021; Roessel *et al*., 2019), especially members of racially and ethnically minoritized groups (Ramjattan, 2022a).

Accent bias may assume heightened importance in newcomers' professional integration (Quinn & Petrick, 1993). Accented L2 speakers often face discrimination as early as in the screening and interview stages (e.g., Ertorer *et al*., 2020; Roessel *et al*., 2019). Anticipating potential bias, they may self-select out of certain job opportunities, fearing that their accent will hinder their hireability (Harrison, 2014). In instances where accented L2 speakers secure employment, they may be subjected to training aimed at "minimizing" their accent (e.g., Thomson, 2014) and helping them speak the "industry standard." Although efforts to eliminate any traces of people's origin are often unsuccessful (Lee *et al*., 2015), the belief that certain professionals should "work on their accent" is widely held by stakeholders in corporate settings, including HR professionals (Chan, 2019; Ramjattan, 2022b; Teló *et al*., 2023; see also Ramjattan, 2021).

Beyond the overarching pattern of downgraded evaluation and reduced employability of accented L2 speakers, accent bias is nuanced (Teló *et al*., 2022). Listeners' evaluations may be influenced by multiple factors, including listeners' prior experience in multilingual and multicultural contexts. For instance, residents of Gainesville (Florida), a location with negative attitudes toward Spanish-speaking bilinguals, exhibited greater bias against accented L2 speakers compared to residents of Montreal (Quebec), where multilingualism is socially embraced (Kutlu

et al., 2022). This finding is particularly noteworthy in light of other work (e.g., Taylor Reid, 2022) that failed to observe differences between listeners from two Canadian cities, Calgary and Montreal, which differ in levels of individual bilingualism but are similar in ethnic diversity.

Accent bias also displays complexity when considering speaker-specific variables such as job type and relative performance level. Accented L2 speakers are perceived as more suitable for lower-prestige jobs, even when their performance is on par with L1-speaking counterparts (Teló *et al*., 2022), and for jobs with low communicative demands such as a cook versus those that require considerable communication skills, including a server or a salesperson (Lindberg and Trofimovich, 2023; Spence *et al*., 2024). In a recent study that manipulated both job prestige and performance level (Teló *et al*., 2022), listeners evaluated L1 English-speaking professionals as more competent than L2 speakers and expressed greater desire to be treated by them than by L2 speakers. However, the same listeners upgraded L2-speaking professionals in their evaluations when these speakers were heard in low-prestige roles and in low-skill scenarios. Thus, listeners may have believed that L2-speaking employees are better suited for low-paying, low-status positions, implying that speaking with an L2 accent might excuse poor professional performance in low-prestige jobs.

The Present Study

Given the importance of advancing the workplace EDI agenda—understood as the establishment of equitable and inclusive work environments where discrimination is absent and every employee can fulfil their potential irrespective of their background or demographic attributes (Ciuk *et al.*, 2022)—it is imperative to investigate how HR students, as future professionals tasked with workplace decision-making, evaluate accented L2 speakers. Despite ample evidence that accented L2 speakers are subject to biased evaluations and unfavorable

decisions, to the best of our knowledge, there is currently no work examining how HR students evaluate such candidates. Given that many HR programs might not focus on linguistic discrimination in the workplace (O'Brien *et al*., 2024), understanding whether HR students demonstrate accent bias is important, especially for the development of future evidence-based pedagogical materials that could be incorporated into EDI curricula. Therefore, we recruited HR students from representative 4-year programs at two major Canadian cities (Calgary, Montreal) to determine how they evaluate the employability of L1- and L2-speaking candidates in mock job interviews.

Because listeners may differ in their familiarity with L2-accented speech and their evaluations might be informed by how accustomed they are to specific accents (Kennedy and Trofimovich, 2008), we compared students' evaluations of L2 speakers from two frequently encountered backgrounds (Arabic and Tagalog). Outside English and French, Arabic is the most widely spoken mother tongue in Montreal, but it ranks sixth in Calgary. By contrast, after English, Tagalog is the second most common language in Calgary, but it is only 15th in Montreal (Statistics Canada, 2021). If listener reactions to accent are context specific (Kutlu *et al*., 2022), rather than reflective of a general bias, then there might be a difference in HR students' assessment of Arabic and Tagalog candidates by location.

We additionally examined students' assessments of candidates at various points in the interview. In previous work, evaluations have been studied at the resume-screening stage (e.g., Oreopoulos, 2011) or in response to audio samples illustrating professional interactions or answers to interview questions (e.g., Kang *et al*., 2023). Although these approaches provide valuable insights, assessments might change over time as more information becomes available (Levon *et al*., 2022; Lindberg and Trofimovich, 2023). Therefore, we asked students to evaluate

the same candidates four times: after reading their resume, after listening to their brief introductory greeting, and after listening to their answers to two interview questions. We also explored students' assessment of job applicants for two comparable jobs (nurse, teacher) to rule out the possibility that listeners might associate certain L2-speaking applicants (e.g., those from the Philippines) with specific jobs (e.g., nursing). Finally, we manipulated the quality of applicants' interview performance (good vs. poor response) as a key methodological control to disentangle students' potential accent bias from their reaction to interview content.

Our study was guided by the following research question: How do HR students from Calgary and Montreal evaluate mock job interview performances of Arabic, English, and Tagalog speakers as a function of job type (nurse vs. teacher), response quality (high vs. low), and timing of assessment (resume vs. introduction vs. content response)? Although HR students might develop some awareness of language and accent bias without any formal training (Trofimovich *et al*., 2024), we anticipated that students would nevertheless favor L1 English speakers over L2-speaking candidates (Spence *et al*., 2024), but that they would provide more favorable assessments to L2 candidates whose accents are more prevalent in their environment (Kennedy and Trofimovich, 2008). We also anticipated that evaluations would not differ by job type, because the two positions have comparable communication demands and are similar in occupational prestige scores and socioeconomic indexes (Hauser and Warren, 1997). On the other hand, we expected low-quality responses to elicit less favorable evaluations than high-quality responses, with L2 applicants especially upgraded in their evaluations when their response quality was low (Teló *et al*., 2022). Finally, with respect to timing, we expected audio-based evaluations (i.e., assessments following introductions and, especially, content responses)

to show the greatest accent bias because audio samples provide clear cues to each candidate's status as an L1 or L2 speaker (Lindberg and Trofimovich, 2023).

Method

Job Advertisements and Resumes

Two job advertisements were created, one for a registered nurse position and another for a math and science teacher position (see Supplementary Information). These materials were based on real job advertisements distributed online, and attended equally to pre-established criteria (e.g., specific job tasks, desired skills, collaborative and highly communicative nature of the work). To provide an evaluative baseline for each candidate (i.e., to determine how they are assessed before students hear their accent), six resumes were created (3 language backgrounds × 2 jobs) in response to the job ads (see Supplementary Information). Identical in layout, resumes provided similar information regarding each candidate's level of education, years of experience, and fit for each job. However, the resumes also presented candidate-specific information related to their background (e.g., name and university attended). Nonetheless, all resumes showed the candidate as an accredited professional in Canada who possessed advanced oral and written proficiency in English and French. No pictures were included in the resumes.

Recordings

Given our goal to investigate evaluations at different points in the interview, two types of scripted scenarios were created for audio recording (introductory statement, question response), where both provided information about each speaker's status as an L1 or L2 speaker, but only the speaker's responses to interview questions included job-relevant information. The applicant's introductory statements included their name and the job to which they were applying. The question responses presented answers to two questions: "How do you respond to constructive

criticism on the job?" and "What personal qualities and experiences make you a good candidate for the job?" The scenarios were adapted from simulated interview recordings by Taylor Reid (2022) to illustrate a high- and a low-quality answer to each question. High-quality responses provided clear and concise answers, included specific examples from the candidate's life, demonstrated enthusiasm, positivity, interest in growth, and appreciation for others. In contrast, low-quality responses were unclear or off-topic, contained impersonal anecdotes, exhibited negativity, lack of enthusiasm and willingness to learn, and spoke poorly about others.

Initially, eight high- and eight low-quality answers were created for each interview question. The scripts were presented to 24 university students from various fields who evaluated their overall quality using a 100-point sliding scale (1 = "very poor response," 100 = "very good response"). The three answers for each question that were rated the highest and the lowest were paired up to compose the final set of scripts, with a reliable difference between the two sets, $t(30) = 12.79$, $p < .001$, $M_{diff} = 49.38$, 95% CI [41.49, 57.36], $d = 4.52$. The scripts were comparable in length ($M = 110.67$ words, $SD = 4.29$) and lexical properties, where each script on average contained 71.17 unique words ($SD = 4.78$) and 98.51% of its lexical content ($SD = 1.38$) was among the 3,000 most commonly used words in English (see Supplementary Information).

Nine speakers (three from each language background, all female) were recorded in individual Zoom meetings with a researcher. The speakers received the scripts before the meeting and were instructed to familiarize themselves with the passages. During the meeting, the speakers were asked to read each script naturally. The two speakers from each background who produced the most natural sounding recordings, as judged by the researchers, were selected to contribute to the final set of recorded scenarios with two introductions (one per job) and one set of answers to each interview question each (one per job), where each response to either question

was unique in its content. The final selection of the recordings included those that had the most suitable pace and most natural flow as evaluated by the researchers. Each speaker received $30 as compensation.

Listeners

To evaluate the applicants, 80 Bachelor of Commerce students majoring in HR- and business-related fields (e.g., HR Management, Marketing, Supply Chain Management) in undergraduate programs in large, English-medium, public universities in Calgary and Montreal (40 per location) were recruited through ads. Inclusion criteria were age (over 18), residence (Calgary or Montreal), and student status (current enrollment in at least one HR management course). On average, students were 21.94 years old (*SD* = 2.74) and had taken 3.64 HR courses (*SD* = 2.42), with the Montreal cohort reporting numerically more completed HR coursework (*M* = 5.02 courses, *SD* = 2.09) than the Calgary cohort (*M* = 2.27 courses, *SD* = 1.78). Fifty-eight students (72.5%) identified as female, while the remaining students identified as male. Students' birthplace included Canada (57.5%) and 20 other countries (42.5%). Similarly, students' L1s included English (41.25%) and 21 other languages (53.75%). Four students (5%) reported L1 Arabic, while no one reported L1 Tagalog. On average, Calgary students knew 2.90 languages (*SD* = 1.20), while Montreal students knew 3.20 languages (*SD* = 1.19). Students reported using English frequently in their daily lives, with Calgary students speaking English 87.58% daily (*SD* = 19.57) and Montreal students speaking English 77.88% daily (*SD* = 19.65). A total of 95.8% of the students self-reported their English proficiency as either "nativelike" or "advanced." All students were fairly familiar with accented English, both in Calgary (*M* = 80.73, *SD* = 21.48) and Montreal (*M* = 76.53, *SD* = 27.48), where 1 meant "not at all familiar" and 100 "very familiar." All students but one reported having normal hearing; because excluding that student's data did

not alter the findings, the entire dataset was considered for analysis. For detailed information on students' background, see Table I of Supplementary Information. Students received $30 for their participation.

Although our listener sample included a substantial number of L2 English speakers, our a priori decision was not to distinguish between listeners based on their L1 or L2 status; we thus refrained from performing separate analyses for L1- and L2-speaking students. From a theoretical perspective, recent meta-analytic evidence indicates that L1- and L2-speaking listeners provide highly comparable impressionistic assessments of speech (Saito, 2021). From a practical perspective, determining a listener's L1 versus L2 status in multilingual and multicultural contexts such as Calgary and Montreal would pose considerable challenges. These cities are home to sizeable proportions of immigrants (approximately 30%), representing over 240 ethnic communities, which exposes residents to a great deal of linguistic diversity. Consequently, even individuals who self-identify as "monolingual native speakers" may be considered multilinguals (Leivada *et al*., 2023). Similar approaches to data analysis have been employed in prior studies investigating accent bias in linguistically and culturally diverse settings (e.g., O'Brien *et al*., 2024; Teló *et al*., 2022; Trofimovich *et al*., 2024). Lastly, this methodological decision follows recent recommendations to treat multilinguals and L2 speakers on equal footing with other language users (e.g., monolinguals, L1 speakers) in research on language attitudes (Truan, 2024). This approach opposes the tendency to view L2 speakers as deficient or otherwise exceptional language users, which perpetuates the marginalization and othering of those speakers. Nevertheless, we accounted for students' linguistic profiles by focusing on other variables capturing their diverse experiences (see below).

Materials and Procedure

Students evaluated three applicants for each job (one from each L1 background) for a total of six applicants, remotely, using a *Qualtrics* web-based interface. Each trial started with the job ad and the candidate's resume, followed by a 100-point sliding scale that prompted students to evaluate the candidate's employability (I would likely hire this candidate… *not at all–very much*). Next, students heard the candidate's introduction and provided the same evaluation. Finally, students heard the candidate's answer to both interview questions, one at a time, providing employability ratings after listening to each answer. An additional 100-point sliding scale presented after each recording targeted the speaker's comprehensibility (I could understand what this person is saying… *not at all–very much*). Students first evaluated the candidates for the nurse position and then proceeded to rating the candidates for the teacher position. The presentation of the candidates within each block was randomized. Before the rating task, students completed a practice trial that presented a candidate for a civil engineer position recorded by a female L1 English speaker. As part of the experimental procedure, students also evaluated the candidates for competence and the extent to which the person was considered an excellent example of someone in that professional occupation. Students were also asked to rank the three candidates applying to each job, and they could provide comments explaining their evaluation. These data, which targeted a conceptually different question and whose treatment would not be possible due to space limitations, are not discussed further.

To account for language background, job type, and response quality, each speaker was presented as a candidate applying to both positions and providing both high- and low-quality answers. Because students had to evaluate each speaker only once, hear a balanced number of speakers from the same language background, evaluate an equal number of high- versus low-

quality answers, and hear the same answer only once, the final set of 36 recordings (2 introductions plus 2 high- and 2 low-quality answers × 6 speakers) was organized into four balanced experimental lists containing six trials as the one described above (three per job), ensuring a balanced distribution of speaker L1, job, and response quality within each list. Across all the lists, students heard each speaker as a candidate applying for each job and illustrating a high- and low-level performance. The 80 HR students were randomly assigned to one of the four versions of the interface, with 10 students per city assigned to each version of the task.

Prior to completing the task, students filled out a consent form and were provided general instructions concerning the structure of the task. Upon completion of the survey, they provided information about their language background and HR experience. This research received ethics approval by the University of Calgary's Conjoint Faculties Research Ethics Board (ID REB21–0116) and by Concordia University's Human Research Ethics Committee (Certificate 30009422).

Data Analysis

Two-way, consistency, average-measures intraclass correlations (ICC) were used to evaluate inter-rater reliability. These indexes, based on the averages of all ratings, estimate the extent to which the HR students (listeners) provided consistent ratings across multiple candidates (speakers). The ICCs were calculated separately across the four survey versions for the Calgary and Montreal cohorts. Overall, students demonstrated highly variable employability ratings after the evaluation of candidates' resumes ($ICC_{Calgary}$ = .16–.40; $ICC_{Montreal}$ = .26–.77) and after listening to their introductory statements ($ICC_{Calgary}$ = .03–.46; $ICC_{Montreal}$ = .37–.63). However, students showed high consistency when rating candidates' employability after listening to their

interview responses ($ICC_{Calgary}$ = .85–.99; $ICC_{Montreal}$ = .92–.96). In the discussion section, we return to this difference in rating consistency as a function of rating time.

We modeled employability ratings through mixed-effects modeling using the lme4 package (Bates *et al.*, 2015) in R version 4.2.2 (R Core Team, 2023), which is an appropriate choice for multi-level data. We ran generalized linear models (using Gaussian distributions) treating candidates' L1 background (Arabic vs. English vs. Tagalog), response quality (high vs. low), timing of employability assessment (four rating episodes), job type (nurse vs. teacher), and location (Calgary vs. Montreal) as fixed-effects predictors, with random intercepts for speakers (6) and students (80). We included six fixed effects as control covariates: (a) students' accent familiarity, (b) students' daily use of spoken English, (c) number of HR courses completed by students, (d) number of languages known by students, (e) students' self-reported gender, and (f) speaker comprehensibility. Control covariates captured additional student- and speaker-specific variance, on the assumptions that employability ratings might differ for students with certain characteristics (e.g., those with greater familiarity with target accents, more HR coursework, or more multilingual experience) and speakers with specific performance attributes (e.g., those who are easier to understand). All continuous control covariates were centered, with the mean of zero and positive and negative values characterizing performance above and below the mean, respectively. Correlations between the response variable and the control covariates are available in Table II of the Supplementary Information.

We evaluated model fit through likelihood ratio tests, comparing a simpler with a more complex model and adopting a more complex model only when it improved fit. Pseudo R^2 values for the models were obtained through the MuMIn package (version 1.47.1; Bartoń, 2020). To explore significant interaction effects, we releveled each relevant model to change the reference

level for a specific categorical variable and refitted the model using the Satterthwaite approximation for degrees of freedom, with *p* values adjusted for multiple comparisons using the Tukey method. The data are publicly available through an Open Science Framework (OSF) study profile (http://doi.org/10.17605/OSF.IO/T9NZX).

Results

Our research question focused on the employability ratings provided by HR students from Calgary and Montreal for job interview performances by L1 Arabic, English, and Tagalog speakers as a function of job type, response quality, and timing of assessment. Figure 1 illustrates these ratings for the three speaker groups, collapsing over location and job type (for simplicity of visual representation, because these variables had no statistically significant impact on the ratings, as described below), separately for high- versus low-quality responses, across the four evaluation timepoints. Globally speaking, the employability ratings for high-quality responses remained high from the initial evaluation of resumes (M = 83.62, SD = 17.25), to the evaluation after the introductory greeting (M = 76.10, SD = 23.81), and through the assessments of the first (M = 81.61, SD = 18.00) and the second (M = 83.79, SD = 18.58) interview responses. In contrast, for low-quality responses, the employability ratings followed a downward trend where evaluations were generally high after the evaluation of resumes (M = 84.27, SD = 17.44) and the introductory greeting (M = 74.73, SD = 25.01) but decreased after the first (M = 31.08, SD = 29.57) and the second (M = 25.67, SD = 28.99) interview responses (see Tables III and IV of the Supplementary Information for full descriptive statistics).

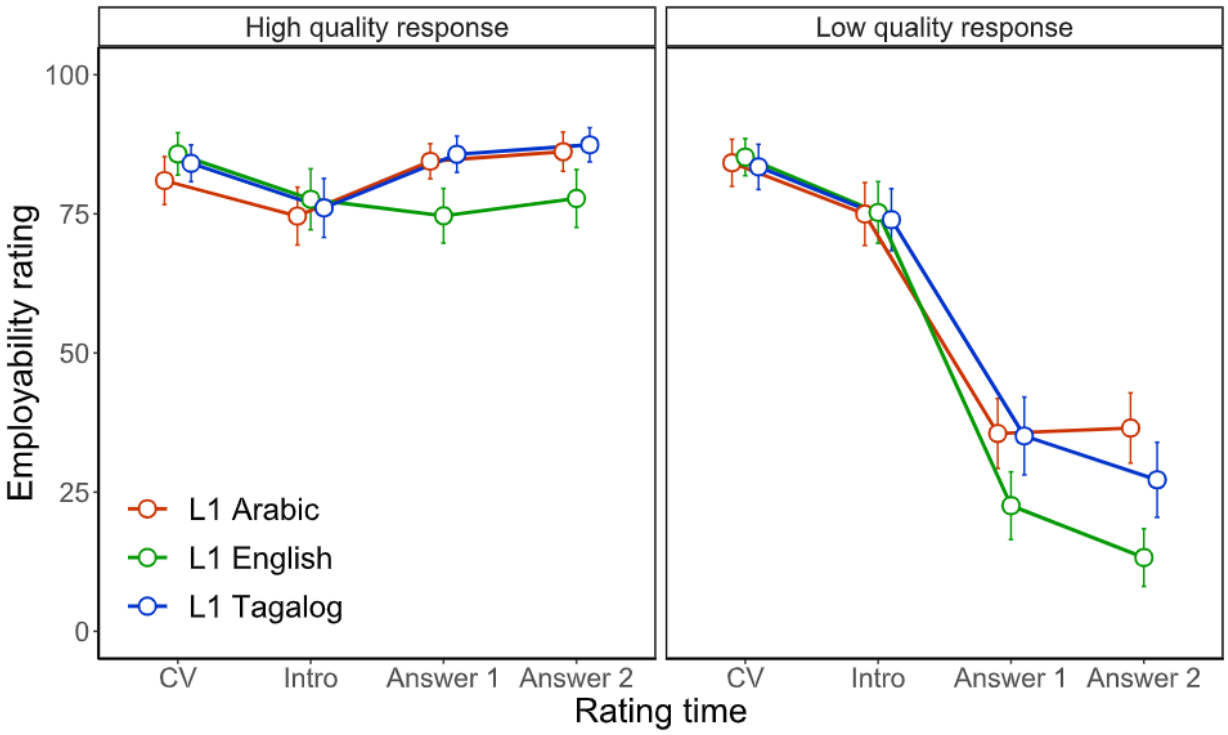

Figure 1. Mean employability ratings and 95% confidence intervals for Arabic, English, and Tagalog speakers in high- versus low-quality interview conditions across four evaluation timepoints: after evaluation of the candidates' resume (labeled "CV" in the figure), after each candidate's introductory statement (Intro), and after each of the two interview responses (Answer 1, Answer 2).

Establishing a Baseline Model

Our initial analysis focused on the global effect of rating time to establish a baseline model for exploring further influences on employability ratings. We fit several models, starting from the intercept model and proceeding to model linear, quadratic, and cubic effects of time. Compared to the initial model, including a linear effect of time significantly improved model fit, $\chi^2(2) = 120.39$, $p < .001$. However, modeling time through quadratic ($p = .998$) or cubic ($p = .996$) functions did not result in improved fit. In addition, modeling by-speaker random slopes for the linear effect of time resulted in singular fit and the model failed to converge, suggesting

overfit. Therefore, we retained only by-speaker and by-student random intercepts. Of the control covariates, only speaker comprehensibility predicted employability ratings, insofar as candidates whose speech was perceived more comprehensible (i.e., easier to understand) received higher employability ratings. Table I summarizes the baseline model (marginal R^2 = .10, conditional R^2 = .19).

Table I. Summary of baseline model

Fixed effects	Estimate	*SE*	95% CI	*t*	*p*
Intercept	73.83	2.50	[68.92, 78.74]	29.51	< .001
Rating time linear	–20.08	1.99	[–23.98, –16.18]	–10.10	< .001
Control covariates					
Accent familiarity	–0.04	1.52	[–3.02, 2.93]	–0.03	.977
English speaking	2.78	1.61	[–0.37, 5.94]	1.73	.084
Number of HR courses	–1.74	1.42	[–4.53, 1.05]	–1.22	.221
Number of languages	2.38	1.44	[–0.44, 5.20]	1.66	.098
Student gender	4.36	3.28	[–2.07, 10.78]	1.33	.184
Speaker comprehensibility	4.21	0.99	[2.27, 6.15]	4.25	< .001
Random intercepts		*SD*			
Student		9.63			
Speaker		3.73			

Note. Table by authors.

Refining the Model

Having established the baseline model of ratings across time, we proceeded to model the remaining effects, retaining only predictors that significantly improved model fit. We started

from speaker-specific effects (speaker L1, response quality), which were of key interest, before modeling occupational and student-specific influences (job type, location), which provided potential nuance to our findings. Using the baseline model as a starting point (Table I), we first added a fixed effect of speaker L1, followed by an interaction of speaker L1 and rating time, which significantly improved model fit, $\chi^2(4) = 15.91$, $p = .003$, implying that speakers from different L1 backgrounds elicited different employability ratings at various rating timepoints. Adding a fixed effect of response quality, followed by an interaction of response quality and rating time, similarly improved model fit, $\chi^2(3) = 1128.70$, $p < .001$, suggesting that responses of high versus low quality elicited different employability ratings at different timepoints. Finally, including an interaction term between speaker L1 and response quality additionally improved model fit, $\chi^2(2) = 8.84$, $p = .012$, indicating that speakers from certain L1 backgrounds were upgraded or downgraded more than others depending on response quality. A three-way interaction of speaker L1, response quality, and rating time had no significant effect on model fit ($p = .239$). We then proceeded to model effects of job type and student location by adding all possible combinations of variables and interactions, including higher-order four- and five-way interactions, but none of these effects improved model fit. Table II summarizes the best-fitting, final model which accounted for 66% of variance in employability ratings (marginal $R^2 = .54$, conditional $R^2 = .66$). Of the control covariates, only speaker comprehensibility significantly predicted employability, where greater speaker comprehensibility was associated with higher employability ratings.

Table II. Summary of the final employability model

Fixed effects	Estimate	*SE*	95% CI	*t*	*p*
Intercept	72.07	2.56	[67.05, 77.08]	28.19	< .001
Speaker L1 × Rating time	–17.13	3.19	[–23.38, –10.88]	–5.37	< .001
Response quality × Rating time	–56.77	2.60	[–61.88, –51.66]	–21.81	< .001
Speaker L1 × Response quality	–7.55	2.60	[–12.66, –2.44]	–2.90	.004
Control covariates					
Accent familiarity	–0.01	1.58	[–3.12, 3.09]	–0.01	.993
English speaking	2.76	1.68	[–0.54, 6.06]	1.64	.101
Number of HR courses	–1.67	1.48	[–4.58, 1.24]	–1.13	.261
Number of languages	2.38	1.50	[–0.56, 5.33]	1.59	.113
Student gender	4.44	3.42	[–2.27, 11.16]	1.30	.195
Speaker comprehensibility	3.59	0.71	[2.21, 4.98]	5.09	< .001
Random intercepts		*SD*			
Student		11.66			
Speaker		0.01			

Note. Speaker L1 = speaker's first language. Table by authors.

Exploring Interaction Effects

We then explored the significant interaction effects obtained in the final employability model. The Speaker L1 × Rating time interaction was driven by students providing higher employability ratings to the L1 Arabic and Tagalog than to L1 English candidates in the last two evaluation episodes (for full comparison details, see Table V of the Supplementary Information). Whereas all speakers (regardless of their L1 background) received similar employability ratings

after the evaluation of their resumes and their introductory statements, L1 Arabic and L1 Tagalog candidates, whose employability ratings did not differ from each other, received higher employability ratings than the L1 English speakers after both responses (by about 12–16 points on a 100-point scale), irrespective of the quality of their answer. Hearing more interview content from the candidates made no difference to students in terms of employability: Whereas the ratings progressively declined across time from the evaluation of resumes to the first interview response (by approximately 23–37 points), they levelled off and remained statistically comparable between the interview responses for each set of candidates (with only 1–3-point difference in ratings).

The significant Response quality × Rating time interaction reflected our built-in manipulation of response quality for the interview responses (for full comparison details, see Table VI of the Supplementary Information). When students evaluated the resumes and the introductory statements, response quality was irrelevant, so all candidates were evaluated similarly. However, when the candidates actually produced high- and low-quality content (in the interview responses), high-quality responses elicited (predictably) higher employability evaluations than low-quality responses (by an average of 50–58 points). In addition, when students heard high-quality responses, candidates' ratings were similar to those they received in the resume stage (with rating differences of 0–7 points between evaluation episodes), meaning that hearing high-quality response content neither diminished nor improved impressions of the candidates based on their resumes. In contrast, hearing low-quality content had a compounding negative effect on evaluations, in the sense that, for low-quality responses, employability declined between each evaluation timepoint, including between the first and the second interview response (with ratings differences of 5–59 points between evaluation episodes).

Finally, the Speaker L1 × Response quality interaction reflected the greater magnitude of between-speaker employability differences for performances featuring low- than high-quality content (for full comparison details, see Table VII of the Supplementary Information). When response quality was high, the L1 Arabic and L1 Tagalog speakers received (significantly) higher employability ratings than the L1 English speakers by about 5–7 points. However, when response quality was low, these differences were magnified, with the L1 Arabic and L1 Tagalog speakers' ratings surpassing those of the L1 English speakers by approximately 9–13 points.

Discussion

Although accented L2 speakers face discrimination in several domains, it remained unclear how HR students—in the absence of formal EDI training on accent and language bias—evaluate L1 and L2 speakers' job interview performances. To address this gap, we examined how HR students in Calgary and Montreal evaluate the employability of L1 Arabic, English, and Tagalog speakers as a function of job type (nurse vs. teacher), response quality (high vs. low), and rating time (resume vs. at various points in the interview). Contrary to our expectations, L1 Arabic and Tagalog speakers (whose evaluations did not differ from each other) generally received higher employability ratings than L1 English speakers, regardless of response quality and irrespective of the amount of interview content (see Figure 1). On the one hand, these results seem to suggest that HR students may be aware of the adversities faced by accented L2 speakers when trying to access the job market in a new country (Trofimovich *et al*., 2024), which resulted in the upgrading of these candidates. This is, undeniably, an encouraging mindset to be observed among future gatekeepers to gainful employment and workplace leaders in advancing the EDI agenda. It is noteworthy that over 50% of the students were themselves L2 English speakers. Although the control covariates capturing students' experiential profiles did not emerge as

significant, their personal experience as L2 speakers may have led to increased empathy toward fellow L2 speakers and influenced their evaluations (Taylor Reid *et al*., 2020). This finding is in line with research utilizing the homophily theory (preferring similar others). For instance, White, L1-speaking hiring decision-makers and those with lower exposure to culture diversity tend to perceive heavier L2 accents as more detrimental to employment decisions (Almeida *et al.*, 2015), whereas more diverse demographic compositions of recruitment teams contribute to more diverse applicant pools (Kazmi *et al*., 2022).

On the other hand, students' preference for L1 Arabic and Tagalog candidates over L1 English candidates was magnified when candidates provided low-quality responses (see Figure 1). Students may have attributed at least part of L2 speakers' poor performance to linguistic limitations that could have prevented them from expressing themselves as they wished (Lindemann and Subtirelu, 2013), acknowledging these limitations through more generous ratings. In support of this interpretation, HR students recently interviewed in Calgary and Montreal expressed a similar sentiment (Trofimovich *et al*., 2024, p. 12), as illustrated in a representative quote from one student: "the mere fact that [L2-speaking job applicants] make a major effort to learn a different language and they can communicate in that language… it's an amazing thing that we should appreciate… and celebrate." Conversely, the students may have evaluated L1 English-speaking candidates particularly harshly, assuming that individuals with full, "native" command of the language should have been (linguistically) capable of answering the interview questions in any desired manner. Regardless of whether students were generous with L2 speakers or demonstrated severity toward L1 speakers, their assessments of candidates may have been grounded on the false premise that a person's accent (pronunciation) reflects their overall language proficiency (Derwing and Munro, 2009). In this study, both L1 and L2 speakers

recorded the same scripts, which were devoid of grammatical errors and featured syntactic and lexical choices commonly employed by L1 speakers. Therefore, all candidates delivered identical answers, with the only manipulated variable being the speaker's accent. In this sense, accent likely functioned for students as a sign of a candidate's linguistic ability, just as stereotypically gay-sounding features in a person's speech might be associated with that speaker's personal traits for listeners (Fasoli *et al*., 2017), with consequences for employability (Fontenele *et al*., 2023).

Considering that the severity of linguistic bias might depend on listeners' age (Taylor Reid *et al*., 2018), it is possible that the students evaluated job candidates' employability differently from older individuals who previously showed a general preference for L1 versus L2 English speakers, including in Calgary (Teló *et al*., 2022). Indeed, our listener sample included young adults (with a mean age of 22 years), whereas Canadian listeners previously demonstrating accent bias were older (with a mean age of around 40 years), so these listeners may have held different conceptions of L2 speakers' employability. Recent evidence from the United States suggests that Generation Z (individuals born after 1997) differs from older generations like Millennials (born between 1981 and 1996) in their perception of English varieties, judging English speakers from North African, Middle Eastern, and South Asian backgrounds more favorably on measures of correctness, pleasantness, and friendliness (Isbell and Crowther, 2023). Assuming that such views could be extended to assessment of employability in interview settings by Generation Z in Canada, it is therefore unsurprising that a young listener sample composed primarily of L2 English speakers such as ours generally found other L2 speakers highly employable.

Age differences and language background alone, however, are unlikely to fully explain why students in this study favorably assessed L2 speakers' employability. Even before hearing the candidates (i.e., when evaluating their resumes), students showed no preference for candidates with English-sounding names and Canadian training and experience, contrary to previous work reporting a bias in favor of candidates with "local" names and cultural capital as well as training and experience in the host country, especially for high-skilled occupations (e.g., Buzdugan and Halli, 2009; Ertorer *et al*., 2020; Oreopoulos, 2011). Upon hearing the candidates' introduction, which confirmed that they were newcomers and L2 speakers, the students once again did not disadvantage L2-speaking candidates in their evaluations. Similarly, even though the students lacked consistency in their evaluations of the candidates early in the interview sequence (after inspecting resumes and hearing personal introductions), their assessments showed high reliability following interview responses, implying a focus on and consideration of job-relevant content.

Despite the general lack of emphasis on accent and language bias in Canadian HR curricula (O'Brien *et al.*, 2024), it is reasonable to believe that our participants received some training on EDI. Our findings seem to reflect, therefore, not only the potential effectiveness of EDI training (Fang *et al*., 2023; Gill *et al*., 2018) but also the trend observed in Canada where diversity management has become a business imperative (Hiranandani, 2012). Canadian employers recognize the strategic value of skilled newcomers (Fang *et al*., 2022, 2023), due in part to the belief that multiculturalism may enhance creativity in the workplace (Leung and Wang, 2015) and that increased diversity and equity may contribute to higher labor productivity and workforce innovation (Armstrong *et al*., 2010). This, in turn, aligns with the flexibility and

open-mindedness shown by Canadian employers and HR students toward both accents and newcomers with respect to hiring practices (Fang *et al*., 2022; Trofimovich *et al*., 2024).

In line with our expectation, compared to low-quality content, high-quality responses elicited higher employability evaluations, regardless of candidate background. This finding supports previous research that similarly manipulated the relative job-specific skill level of the speaker (Teló *et al*., 2022, 2024). Interestingly, however, low-quality responses were particularly detrimental to candidates' employability over time, whereas high-quality responses maintained consistent evaluations throughout the interview sequence (see Figure 1), suggesting that high- and low-quality responses were subject to different types of impression formation. High-quality interview content appeared to be cumulatively redundant, in the sense that the answers to both interview questions were of similar (high) quality and thus provided no additional information about the candidate (Alves and Mata, 2019). Conversely, low-quality interview content was likely encoded incrementally following the summation rule (Betsch *et al*., 2001), where students aggregated successive negative observations to form their assessment of the candidate. This distinction in how students weighed high- versus low-quality content toward their evaluations can also be interpreted through the lens of the negativity bias perspective (Rozin and Royzman, 2001), where people tend to attend to and recall negative information more heavily than positive information (Unkelbach *et al*., 2020). Regardless of its origin, this difference in students' reactions to high- versus low-quality interview content over time applied to all speakers irrespective of their language background, implying that these evaluative forces are independent of a candidate's language background.

Also consistent with our predictions, job type (nurse vs. teacher) and rater location (Calgary vs. Montreal) did not interact with employability when entered into statistical models.

Notwithstanding previous research documenting biases against actual L2-speaking nurses and teachers or research participants assuming these roles (e.g., Kang *et al*., 2023; Neiterman and Bourgeault, 2015), we expected students to provide comparable evaluations for candidates applying to those positions given that they have similar communication demands and are of similar prestige. Moreover, drawing from prior research on listener evaluations of L2 speech in Calgary and Montreal (Taylor Reid, 2022) and considering the comparable levels of ethnic diversity in both cities (Statistics Canada, 2021), we anticipated that HR students from the two cities would provide similar evaluations. Additionally, students did not seem to favor the L2 English variety that was more prevalent in their location (Tagalog in Calgary, Arabic in Montreal), so students' employability ratings may not have been driven by the frequency of exposure to specific L2 accents. The lack of significant effects of student location on employability as well as students' general preference for L2 over L1 English speakers thus support the positive influence of ethnically and linguistically diverse settings on assessments of accented L2 speech (Kutlu *et al*., 2022).

Finally, in keeping with our goal to investigate general patterns in HR students' employability evaluations of L1 and L2 English-speaking candidates, our statistical model included five control covariates pertaining to students' background characteristics, in addition to the control covariate capturing speaker comprehensibility. The only variable that emerged as significant was speaker comprehensibility, such that candidates who were easier to understand were considered to be more employable. While we are not aware of any studies that investigated the role of various listener background variables in employability decisions by HR students, the general preference for speakers whose pronunciation is easier to understand—including among HR professionals—is well established. Briefly, when listeners find a speaker difficult to

understand, as they struggle to decode the speaker's message, they might downgrade the speaker in their evaluations (Dragojevic and Giles, 2016; Teló *et al*., 2022, 2023).

Limitations, Future Research, and Implications

Despite their promising outlook, our findings need to be interpreted in light of several limitations. First, the generalizability of our results may be constrained by the specific jobs examined. While those occupations are typically considered of average prestige in Canada, it is possible that L1 English-speaking candidates would be preferred over accented L2 speakers were we targeting high-prestige jobs (Teló *et al*., 2022). Similarly, nurses and teachers employed in diverse settings might be expected to interact with similarly diverse groups of individuals such as patients and schoolchildren, and these assumptions might influence decision-makers' evaluations (Almeida *et al.*, 2015). Second, the low consistency observed in employability ratings after the evaluation of candidates' resumes and after listening to their introductory statements calls for caution in interpreting these evaluations. For resumes, low consistency in evaluations likely reflects students' difficulty making a reliable assessment based on limited information (i.e., resume only). For introductory statements, students' highly variable assessments likely stem from the low ecological validity of this rating episode, with only the applicant's name and the position to which they were applying available to students (e.g., "*Hi! My name is Zein Ahmad and I'm applying for the registered nurse position*"), which may understandably lead to highly variable evaluations (i.e., some students may believe that it is difficult to evaluate employability based on introductions alone and assign low values, whereas others may view the introductions as non-risky and assign high values). In future work, if researchers wish to evaluate candidates on multiple occasions, they could include samples that provide similar amounts of information but differ in job-specific content (e.g., similar-length recordings featuring candidates' pre-

interview small talk vs. responses to interview questions). Third, considering the influence that EDI training may have had on HR students' evaluations of L1- and L2-speaking candidates, mixed methods studies that include in-depth analysis of the nature and scope of EDI training received by the participants are welcome. Fourth, since none of the control covariate pertaining to students' background emerged as significant, we invite researchers to further investigate these and other variables as they might provide insights into HR students' behavior. Lastly, because our data were collected in an experimental setting, it is important to acknowledge that students may have been aware that their attitudes toward L2 speakers and newcomers were being tested, and as HR students, may have consciously chosen to demonstrate high levels of acceptance.

As the first study investigating HR students' evaluations of L1 and L2 speakers in simulated job interviews, our work offers several avenues for future research with practical implications for HR practice. The overall positive reception of L1 Arabic and Tagalog job applicants likely reflects individual and combined influences of the diverse, multicultural settings where the study was conducted, the EDI initiatives with which students might have engaged during their training, and the evolving mandate for diversity management in Canada. This, in turn, raises questions about the relationship between a student's location and the nature of the EDI training they receive, with implications for their decision-making. For example, students residing in multicultural and multilingual locales, where they frequently interact with L2 speakers, may require less emphasis on language and accent bias in their training. Conversely, in less diverse locations, EDI training and initiatives may need to place a greater focus on exposure to and interactions with a range of L2 speakers. Future research could examine the impact of tailoring EDI instruction to account for the specific audience, their experiences, and the context where they operate with the goal of determining the most beneficial outcomes for HR students in

terms of helping them redefine how they think about linguistic diversity and the decisions they make as a result—ultimately contributing to the advancement of the EDI agenda.

References

Almeida, S., Fernando, M., Hannif, Z. and Dharmage, S.C. (2015), "Fitting the mould: the role of employer perceptions in immigrant recruitment decision-making", *The International Journal of Human Resource Management*, Vol. 26 No. 22, pp.2811–2832. https://doi.org/10.1080/09585192.2014.1003087

Alves, H. and Mata, A. (2019), "The redundancy in cumulative information and how it biases impressions", *Journal of Personality and Social Psychology*, Vol. 117 No. 6, pp.1035–1060. http://doi.org/10.1037/pspa0000169

Armstrong, C., Flood, P.C., Guthrie, J.P., Liu, W., MacCurtain, S. and Mkamwa, T. (2010), "The impact of diversity and equality management on firm performance: beyond high performance work systems", *Human Resource Management*, Vol. 49 No. 6, pp.977–998. https://doi.org/10.1002/hrm.20391

Baquiran, C.L.C. and Nicoladis, E. (2020), "A doctor's foreign accent affects perceptions of competence", *Health Communication*, Vol. 35 No. 6, pp.726–730. https://doi.org/10.1080/10410236.2019.1584779

Bartoń, K. (2023), "*MuMIn: multi-model inference*. R package version 1.47.1, available at: https://CRAN.R-project.org/package=MuMIn

Bates, D., Mächler, M., Bolker, B. and Walker, S. (2015), "Fitting linear mixed-effects models using lme4", *Journal of Statistical Software*, Vol. 67 No. 1, pp.1–48. https://doi.org/10.18637/jss.v067.i01

Betsch, T., Plessner, H., Schwieren, C. and Gütig, R. (2001), "I like it but I don't know why: a value-account approach to implicit attitude formation", *Personality and Social*

Psychology Bulletin, Vol. 27 No. 2, pp.242–253. https://doi.org/10.1177/0146167201272009

Buzdugan, R. and Halli, S.S. (2009), "Labor market experiences of Canadian immigrants with focus on foreign education and experience", *International Migration Review*, Vol. 43 No. 2, pp.366–386. https://doi.org/10.1111/j.1747-7379.2009.00768.x

Chan, J.Y.H. (2019), "The choice of English pronunciation goals: different views, experiences and concerns of students, teachers and professionals", *Asian Englishes*, Vol. 19 No. 3, pp.264–284. ttps://doi.org/10.1080/13488678.2018.1482436

Ciuk, S., Śliwa, M. and Harzing, A.-W. (2022), "Implementing the equality, diversity, and inclusion agenda in multinational companies: a framework for the management of (linguistic) diversity", *Human Resource Management Journal*. https://doi.org/10.1111/1748-8583.12487

Comer, D.R., Lenaghan, J.A. and Motro, D. (2023), "Seeing past different signals in the job interview: information improves ratings of candidates on the autism spectrum", *Equality, Diversity and Inclusion*, Vol. 42 No. 7, pp.872–888. https://doi.org/10.1108/EDI-12-2022-0334

Derwing, T.M. and Munro, M.J. (2009), "Putting accent in its place: rethinking obstacles to communication", *Language Teaching*, Vol. 42 No. 4, pp.476–490. https://doi.org/10.1017/S026144480800551X

Dragojevic, M. and Giles, H. (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*, Vol. 42 No. 3, pp.396–420. https://doi.org/10.1111/hcre.12079

Ertorer, S.E., Long, J., Fellin, M. and Esses, V.M. (2020), "Immigrant perceptions of integration in the Canadian workplace", *Equality, Diversity and Inclusion*, Vol. 41 No. 7, pp.1091–1111. https://doi.org/10.1108/EDI-02-2019-0086

Fang, T., Xiao, N., Zhu, J. and Hartley, J. (2022), "Employer attitudes and the hiring of immigrants and international students: evidence from a survey of employers in Atlantic Canada", *Canadian Public Policy*, Vol. 48 No. 1, pp.17–37. https://doi.org/10.3138/cpp.2022-007

Fang, T., Zhang, T. and Hartly, J. (2023), "Examining determinants of employers' attitudes toward hiring immigrant workers: evidence from an employer survey", *Journal of Immigrant & Refugee Studies*, https://doi.org/10.1080/15562948.2023.2219634

Fasoli, F., Maass, A., Paladino, M.P. and Sulpizio, S. (2017), "Gay- and lesbian-sounding auditory cues elicit stereotyping and discrimination", *Achieves of Sexual Behavior*, Vol. 46 No. 5, pp.1261–1277. https://doi.org/10.1007/s10508-017-0962-0

Fontenele, A.B.G., Souza, L.E.C. and Fasoli, F. (2023), "Who does discriminate against gay-sounding speakers? The role of prejudice on voice-based hiring decisions in Brazil", *Journal of Language and Social Psychology*, Vol. 42 No. 1, pp.119–130. https://doi.org/10.1177/0261927X221077243

Gill, G.K., McNally, M.J. and Berman, V. (2018), "Effective diversity, equity, and inclusion practices", *Healthcare Management Forum*, Vol. 31 No. 5, pp.196–199. https://doi.org/10.1177/084047041877378

Harrison, G. (2014), "Accent and 'othering' in the workplace", Levis J.M and Moyer A. (Ed.s), *Social dynamics in second language accent*, De Gruyter, Berlin, pp.255–272. https://doi.org/10.1515/9781614511762.255

Hauser, R.M. and Warren, J.R. (1997), "Socioeconomic indexes for occupations: a review, update, and critique", *Sociological Methodology*, Vol. 27 No. 1, pp.177–298. https://doi.org/10.1111/1467-9531.271028

Hiranandani, V. (2012), "Diversity management in the Canadian workplace: towards an antiracism approach", *Urban Studies Research*, Vol. 2012, pp.1–13. https://doi.org/10.1155/2012/385806

Isbell, D.R. and Crowther, D. (2023), "Revisiting US undergraduate perceptions of non-native English varieties: from Millennials to Generation Z", *Applied Linguistics*, Vol. 44 No. 2, pp.287–311. https://doi.org/10.1093/applin/amac048

Kang, O., Yaw, K. and Kostromitina, M. (2023), "The effects of situational contexts and occupational roles on listeners' judgements on accented speech", *Psychology of Language and Communication*, Vol. 27 No. 1, pp.1–22. http://doi.org/10.58734/plc-2023-0001

Kazmi, M.A., Spitzmueller, C., Yu, J., Madera, J.M., Tsao, A.S., Dawson, J.F. and Pavlidis, I. (2022), "Search committee diversity and applicant pool representation of women and underrepresented minorities: a quasi-experimental field study", *Journal of Applied Psychology*, Vol. 107 No. 8, pp.1414–1427. http://doi.org/10.1037/APL0000725

Kennedy, S. and Trofimovich, P. (2008), "Intelligibility, comprehensibility, and accentedness of L2 speech: the role of listener experience and semantic context", *Canadian Modern Language Review*, Vol. 64 No. 3, pp.459–490. https://doi.org/10.3138/cmlr.64.3.459

Kutlu, E., Tiv, M., Wulff, S. and Titone, D. (2022), "Does race impact speech perception? An account of accented speech in two different multilingual locales", *Cognitive Research:*

Principles and Implications, Vol. 7 No. 1, pp.1–16. https://doi.org/10.1186/s41235-022-00354-0

Lee, J., Jang, J. and Plonsky, L. (2015), "The effectiveness of second language pronunciation instruction: a meta-analysis", *Applied Linguistics*, Vol. 36 No. 3, pp.345–366. https://doi.org/10.1093/applin/amu040

Leivada, E., Rodríguez-Ordóñez, I., Couto, M.C.P. and Perpiñán, S. (2023), "Bilingualism with minority languages: Why searching for unicorn language users does not move us forward", *Applied Psycholinguistics*, Vol. 44 No. 3, pp.384–399. https://doi.org/10.1017/S0142716423000036

Leung, K. and Wang, J. (2015), "Social processes and team creativity in multicultural teams: a socio-technical framework", *Journal of Organizational Behavior*, Vol. 36 No. 7, pp.1008–1025. https://doi.org/10.1002/job.2021

Lev-Ari, S. and Keysar, B. (2010), "Why don't we believe non-native speakers? The influence of accent on credibility", *Journal of Experimental Social Psychology*, Vol. 46 No. 6, pp.1093–1096. https://doi.org/10.1016/j.jesp.2010.05.025

Levon, E., Sharma, D. and Ye, Y. (2022), "Dynamic sociolinguistic processing: real-time changes in judgments of speaker competence", *Language*, Vol. 98 No.4, pp.749–774. http://doi.org/10.1353/lan.2022.0020

Lindberg, R. and Trofimovich, P. (2023), "When accent does not match expectations: a dynamic perspective of L2 speaker evaluations in a French interview context", *Canadian Journal of Applied Linguistics*, Vol. 26 No. 2, pp.1–31. https://doi.org/10.37213/cjal.2023.33052

Lindemann, S. and Subtirelu, N. (2013), "Reliably biased: the role of listener expectation in the perception of second language speech", *Language Learning*, Vol. 63 No. 3, pp.567–594. https://doi.org/10.1111/lang.12014

Näre, L. and Cleland Silva, T. (2021), "The global bases of inequality regime: the case of international nurse recruitment", *Equality, Diversity and Inclusion*, Vol. 40 No. 5, pp.510–524. https://doi.org/10.1108/EDI-02-2020-0039

Neiterman E. and Bourgeault, I.L. (2015), "The shield of professional status: comparing internationally educated nurses' and international medical graduates' experiences of discrimination", *Health*, Vol. 19 No. 6, pp.615–634. http://doi.org/10.1177/1363459314567788

O'Brien, M.G., Le, T.-N.N., Bodea, A., Trofimovich, P. and Shimada, M. (2024), "Accent bias training in undergraduate human resource management education", *Business and Professional Communication Quarterly*. https://doi.org/10.1177/23294906231226212

Oreopoulos, P. (2011), "Why do skilled immigrants struggle in the labor market? A field experiment with thirteen thousand resumes", *American Economic Journal: Economic Policy*, Vol. 3 No. 4, pp.148–171. http://doi.org/10.1257/pol.3.4.148

Quinn, J.F. and Petrick, J.A. (1993), "Emerging strategic human resource challenges in managing accent discrimination and ethnic diversity", *Applied H.R.M Research*, Vol. 4 No. 2, pp.79–93.

R Core Team (2023), "*R: a language and environment for statistical computing*", R Foundation for Statistical Computing, available at: https://www.R-project.org

Ramjattan, V.A. (2021), “The transracial aesthetic labour of an international teaching assistant”, *Equality, Diversity and Inclusion*, Vol. 40 No. 8, pp.973–985. https://doi.org/10.1108/EDI-12-2020-0365

Ramjattan, V.A. (2022a), “Accenting racism in labour migration”, *Annual Review of Applied Linguistics*, Vol. 42 No. 1, pp.87–92. http://doi.org/10.1017/S0267190521000143

Ramjattan, V.A. (2022b), “Racializing the problem and solution to foreign accent in business”, *Applied Linguistics Review*, Vol. 13 No. 4, pp.527–544. https://doi.org/10.1515/applirev-2019-0058

Roessel, J., Schoel, C., Zimmermann, R. and Stahlberg, D. (2019), “Shedding new light on the evaluation of accented speakers: basic mechanisms behind nonnative listeners’ evaluations of nonnative accented job candidates”, *Journal of Language and Social Psychology*, Vol. 38 No. 1, pp.3–32. https://doi.org/10.1177/0261927X17747904

Rozin, P. and Royzan, E.B. (2001), “Negativity bias, negativity dominance, and contagion”, *Personality and Social Psychology Review*, Vol. 5 No. 4, pp.296–320. https://doi.org/10.1207/S15327957PSPR0504_2

Saito, K. (2021), “What characterizes comprehensible and native-like pronunciation among English-as-second-language speakers? Meta-analyses of phonological, rater, and instructional factors”, *TESOL Quarterly*, Vol. 55 No. 3, pp.866–900. https://doi.org/10.1002/tesq.3027

Spence, J.L., Hornsey, M.J., Stephensen, E.M. and Imuta, K. (2024), “Is your accent right for the job? A meta-analysis on accent bias in hiring decisions”, *Personality and Social Psychology Bulletin*, Vol. 50 No. 3, pp.371–386. https://doi.org/10.1177/01461672221130595

Statistics Canada (2021), "*2021 census of population*", available at: https://www12.statcan.gc.ca/census-recensement/2021/dp-pd/prof/index.cfm?Lang=E#shr-pg0 (accessed 7 October 2023)

Sue, D.W., Capodilupo, C.M., Torino, G.C., Bucceri, J.M., Holder, A.M.B., Nadal, K.L. and Esquilin, M. (2007), "Racial microaggressions in everyday life: implications for clinical practice", *American Psychologist*, Vol. 62 No. 4, pp.271–286. https://doi.org/10.1037/0003-066X.62.4.271

Taylor Reid, K. (2022), *Exploring social attitudes toward second language speakers of English across Canada*. PhD thesis. Concordia University. https://spectrum.library.concordia.ca/id/eprint/991004/1/Taylor_PhD_F2022.pdf

Taylor Reid, K., O'Brien, M.G., Trofimovich, P. and Bajt, A.M. (2020), "Testing the malleability of teachers' judgments of second language speech", *Journal of Second Language Pronunciation*, Vol. 6 No. 2, pp.236–264. https://doi.org/10.1075/jslp.19015.tay

Taylor Reid, K., Trofimovich, P., and O'Brien, M.G. (2018), "Social attitudes and speech ratings: effects of positive and negative bias on multiage listeners' judgments of second language speech", *Studies in Second Language Acquisition*, Vol. 41 No. 2, pp.419–442. https://doi.org/10.1017/S0272263118000244

Teló, C., O'Brien, M.G., Trofimovich, P., Bodea, A., Le, T.-N.N. and Shimada, M. (2023), "How HR professionals respond to second language accents", *Talent Canada*, 20 January, https://www.talentcanada.ca/how-hr-professionals-respond-to-second-language-accents/

Teló, C., Silveira, R., Marcelino, A.F.B. and O'Brien, M.G. (2024), "Accent bias in professional evaluations: a conceptual replication study in Brazil", *Applied Linguistics*, pp. 1–20. http://doi.org/10.1093/applin/amae042.

Teló, C., Trofimovich, P. and O'Brien, M.G. (2022), "Disentangling professional competence and foreign accent", *Journal of Second Language Pronunciation*, Vol. 8 No.3, pp.413–443. https://doi.org/10.1075/jslp.22013.tel

Thomson, R.I. (2014), "Myth 6: accent reduction and pronunciation instruction are the same thing", Grant, L (Ed.), *Pronunciation myths: applying second language research to classroom teaching*, University of Michigan Press, Ann Arbor, MI, pp.160–187. http://doi.org/10.3998/mpub.4584330

Trofimovich, P., Bodea, A., Le, T.-N.N., O'Brien, M.G., Shimada, M. and Teló, C. (2024), "What do students in human resource management know about accent bias?" *Language Awareness*, Vol. 33 No. 3, pp. 447–467. https://doi.org/10.1080/09658416.2023.2179630

Truan, N. (2024), "Whose language counts? Native speakerism and monolingual bias in language ideological research: challenges and directions for future research" *European Journal of Applied Linguistics*. https://doi.org/10.1515/eujal-2024-0006

Unkelbach, C., Alves, H. and Koch, A. (2020), "Negativity bias, positivity bias, and valence asymmetries: explaining the differential processing of positive and negative information", Gawronski, B. (Ed.), *Advances in experimental social psychology*, Elsevier, New York, NY, pp.115–187. https://doi.org/10.1016/bs.aesp.2020.04.005

[1] In this article, we use the term "accented L2 speakers" to refer to the individuals who learned their L2 as an additional language (i.e., after the acquisition of their first language), who are typically born and raised outside the country where they use their L2, and whose L2 speech is marked by perceptible pronunciation features that mark them as outsiders. We purposely avoid

the use of the terms “foreign” or “non-native” to qualify L2 accents or speakers, or to juxtapose them with “native” accents or speakers, because these terms tend to cue deficiency-based views of language and language speakers, in the sense that “foreign” and “non-native” accents are often understood as something to be rid of in favor of the elusive “native” standard. More importantly, our use of “accented L2 speakers” to refer to speakers whose L2 pronunciation is influenced by their previously learned languages is not tied to language proficiency; indeed, many accented L2 speakers speak fluently, use appropriate vocabulary and grammar, and produce pragmatically and contextually relevant discourse.