

Exploring social attitudes toward second language speakers of English across Canada

Kimberly Taylor

A Thesis

In the Department

of

Education

Presented in Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy (Education) at

Concordia University

Montréal, Québec, Canada

July 2022

© Kimberly Taylor, 2022

CONCORDIA UNIVERSITY
SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Kimberly Taylor

Entitled: *Exploring social attitudes toward second language speakers of English across Canada*

and submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY (Education)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

_____ Chair

Dr. Diane Pesco

_____ External Examiner

Dr. Suzie Beaulieu

_____ Examiner

Dr. Kim McDonough

_____ Examiner

Dr. Walcir Cardoso

_____ Examiner

Dr. Sarita Kennedy

_____ Examiner

Dr. Mary O'Brien

_____ Thesis Supervisor

Dr. Pavel Trofimovich

Approved by:

Dr. Saul Carliner, Department Chair _____

Dr. Pascale Sicotte, Dean of Faculty _____

August 24, 2022

ABSTRACT

Exploring social attitudes toward second language speakers of English across Canada

Kimberly Taylor, PhD

Concordia University, 2022

Nearly a fifth of Canada's population is represented by people from other countries, most of whom speak languages other than English or French. Though there has been some exploration of social attitudes toward these ethnolinguistic groups, there have been no multi-city investigations of social attitudes, particularly toward second language (L2) speakers, in major cities where immigrant populations are most concentrated. Furthermore, there have been few studies that present speech samples alongside images suggesting speaker ethnicity and/or religious affiliation. Therefore, this dissertation explores ratings of L2 speakers across multiple dimensions and considers the role of social attitudes and social network exposure in formation of those judgments.

Study 1 explored how residents of Calgary and Montreal judge the comprehensibility and accentedness of L2 speech in audio-only and audiovisual conditions and whether those judgments are associated with residents' overall social attitudes toward immigrants. There were no context or image effects, but differences emerged among ratings of certain language groups, as well as raters' general attitudes toward immigrants. Ultimately, raters' attitudes were not associated with their ratings of L2 speech.

Study 2 explored how native-born residents of Canada judge L2 speakers' intelligence, friendliness, and trustworthiness and investigated how those judgments might be related to residents' general social attitudes toward immigrants. No significant differences emerged between speech samples presented as audio-only versus with a nonreligious or religious image.

However, there was a clear hierarchy in how specific language groups were evaluated. Social attitudes questionnaire responses revealed generally positive attitudes toward immigrants. Ultimately, those attitudes had weak relationships with rater judgments of L2 speakers' intelligence, friendliness, and trustworthiness.

Study 3 explored how native-born residents of Montreal and Calgary compared in judgments of L2 speaker citizenship and how those judgments might be related to raters' L2 social networks. No differences based on image condition or context surfaced in judgments of L2 speaker citizenship. However, specific language groups differed in how their citizenship status was perceived. There were also between-context differences in native-born residents' interactions with L2 speakers. Ultimately, social network characteristics had no influence on citizenship ratings in either context.

Acknowledgments

Completion of this monumental task would not have been possible without an army of support. First, I would like to thank my brilliant supervisor, Dr. Pavel Trofimovich. Words cannot adequately express my gratitude for his mentorship and kindness while I pursued two degrees and weathered countless life upheavals. Through it all, he has maintained unwavering confidence in me, even when I lacked confidence in myself. We share the same sense of optimism, excitement about new project ideas, and love for really good writing. He is, indeed, a kindred spirit. I have often thought about how rare it must be to find a supervisor who takes every opportunity to shift the spotlight away from himself in favour of his students' recognition, and I feel honoured to have benefitted from his knowledge for so long.

Through our work together, Pavel has introduced me to a second talented researcher, Dr. Mary O'Brien, who has been my unofficial mentor and sounding board throughout my doctoral journey. The three of us have collaborated on countless projects of which I am immensely proud. I am so grateful for the time, energy, and knowledge she has shared with me. I am also deeply grateful to my advisory committee members, Dr. Kim McDonough, Dr. Walcir Cardoso, and Dr. Sarita Kennedy for sharing their expertise, humour, and honesty and for helping me grow considerably as a researcher. I extend further thanks to Dr. Suzie Beaulieu for agreeing to join the committee as an external examiner and to Dr. Diane Pesco for agreeing to chair my defense.

There have been many Concordians, past and present, who have held me up or helped me out at some point in my academic journey, so many that I could not possibly name them all here, but I am particularly grateful to Jennica Grimshaw, Lauren Strachan, June Ruivivar, Aki Tsunemoto, Annie Bergeron, Clinton Hendry, and Jenn Foote for being part of my story. My decision to hop the border so many years ago in search of a new adventure was a good one.

Finally, I want to thank my family and friends for being my constant cheerleaders and for having such faith in my abilities. My children, Alex and Macy, have been remarkably patient throughout this journey and have always shown great pride in their mom. Little do they know how much I look up to them. My love and deep gratitude also go to my mom, my dad, my sister Su, the rest of my spectacular family, and my dear friends, especially Anna, Janine, and Maeghan. This wonderful collection of folks has always made me feel like I could achieve anything. As my cutie-pie-of-a-mother often says, “I am a billionaire.”

This research was supported by the Social Sciences and Humanities Research Council.



Dedication

To my children, Alex and Macy, two extraordinary humans

Contribution of Authors

The three studies in this dissertation are co-authored with my supervisor, Dr. Pavel Trofimovich.

Studies 1 and 2 have been presented at international conferences: The American Association of Applied Linguistics (AAAL) and Pronunciation in Second Language Learning and Teaching (PSLLT), respectively, and all three manuscripts are being prepared for journal submission.

Outlined below are individual contributions:

Research design: Kym and Pavel

Recruitment and data collection: Kym

Materials and survey creation: Kym with feedback from Pavel

Data analysis: Kym with advice from Pavel

Manuscript writing: Kym with feedback from Pavel

Funding: Drs. Pavel Trofimovich and Mary O'Brien, SSHRC Insight Development Grant

Table of Contents

List of Figures	xiii
List of Tables	xiv
Definition of Key Terms	xvi
Chapter 1: General Introduction	1
Social Bias and Language	3
Language-Based Bias in Canada	6
The Dissertation Studies	9
Chapter 2: Study 1	13
The Roots of Group Conflict	13
Exploring Ethnic and Linguistic Biases	15
The Current Study	19
Method	23
Raters	23
Materials	24
Speech Samples and Photo Guises	24
Speech Ratings	25
Social Attitudes Questionnaire	27
Procedure	28
Data Analysis	28
Results	30
Speech Ratings	30
Attitudes About Immigration	36
Relationship Between Speech Ratings and Attitudes About Immigration	38
Discussion	40
Ratings of Comprehensibility and Accentedness	40
Social Attitudes About Immigrants	44
Relationship Between Social Attitudes and Speech Ratings	45

Calgary – Montreal Comparison: What Have We Learned?	47
Implications and Future Research.....	49
Limitations	51
Conclusion	52
Connecting Study 1 and Study 2.....	53
Chapter 3: Study 2.....	54
Sources of Bias	54
Status and Solidarity Judgments	56
The Current Study.....	59
Method	64
Raters	64
Materials	65
Speech Samples and Photo Guises	65
Speech Ratings.....	66
Social Attitudes Questionnaire	66
Procedure	68
Data Analysis	69
Results.....	70
Status and Solidarity Ratings	70
Attitudes About Immigration and Relationship to Speech Ratings	78
Discussion	80
Ratings of Intelligence, Friendliness, and Trustworthiness	81
Social Attitudes About Immigrants	86
Implications and Future Research.....	88
Limitations	90
Conclusion	91
Connecting Study 2 and Study 3.....	91
Chapter 4: Study 3.....	93

Perceptions and Social Bias	94
Language-Based Biases	95
The Role of Social Networks	97
The Current Study	100
Method	104
Raters	104
Materials	105
Speech Samples and Photo Guises.....	105
Speech Ratings.....	106
Social Network Survey.....	107
Procedure	108
Data Analysis	108
Results.....	110
Citizenship Labels.....	110
Social Network Surveys.....	113
Social Network Characteristics and Relationship to Citizenship Labeling	114
Discussion.....	115
Citizenship Labels.....	116
Language-Based Citizenship Hierarchy	119
Social Networks	122
Implications.....	123
Limitations and Future Research	124
Conclusion	126
Chapter 5: General Discussion and Conclusion.....	127
Overview of Studies.....	127
Key Findings.....	127
Overall Implications.....	130
Understanding Attitudes in Canada	130

Theoretical Implications	132
Societal Implications.....	133
Limitations and Suggestions for Future Research	136
Conclusion	138
References.....	139
Appendix A: Sample Image Guises for Speech Rating Task	175
Appendix B: Image Pre-Rating Task.....	176
Appendix C: Social Attitudes Questionnaire.....	178
Appendix D: Full Statistical Output for Mixed-Effects Models in Study 1	180
Appendix E: Full Statistical Output for Mixed-Effects Models in Study 2.....	194
Appendix F: Social Network Survey	209
Appendix G: Full Statistical Output for Mixed-Effects Models in Study 3	212

List of Figures

Figure 2.1 <i>Comprehensibility Ratings by Speaker Group and Location</i>	35
Figure 2.2 <i>Accentedness Ratings by Speaker Group and Location</i>	36
Figure 3.1 <i>Intelligence Ratings by Speaker Group</i>	76
Figure 3.2 <i>Friendliness Ratings by Speaker Group</i>	77
Figure 3.3 <i>Trustworthiness Ratings by Speaker Group</i>	78
Figure 4.1 <i>Citizenship Labeling by Speaker Group</i>	113

List of Tables

Table 2.1 Means (Standard Deviations) for Comprehensibility and Accentedness Ratings by Speaker Group, Condition, and Location.....	31
Table 2.2 Summary of Mixed-Effects Models for Comprehensibility Ratings in the Voice–Image Match and Mismatch Conditions by Location and Speaker Group.....	32
Table 2.3 Summary of Mixed-Effects Models for Accentedness Ratings in the Voice–Image Match and Mismatch Conditions by Location and Speaker Group.....	33
Table 2.4 Means (Standard Deviations) for Raters’ Responses to the Social Attitudes Questionnaire by Location.....	37
Table 2.5 Summary of Mixed-Effects Models for Raters’ Responses to the Social Attitudes Questionnaire.....	38
Table 2.6 Pearson Correlations Between Raters’ Speech Ratings and Their Responses to the Social Attitudes Questionnaire.....	39
Table 3.1 Means (Standard Deviations) for Intelligence, Friendliness, and Trustworthiness Ratings by Speaker Group and Condition.....	72
Table 3.2 Summary of Mixed-Effects Models for Intelligence Ratings in the Nonreligious and Religious Conditions by Speaker Group.....	73
Table 3.3 Summary of Mixed-Effects Models for Friendliness Ratings in the Nonreligious and Religious Conditions by Speaker Group.....	74
Table 3.4 Summary of Mixed-Effects Models for Trustworthiness Ratings in the Nonreligious and Religious Conditions by Speaker Group.....	75
Table 3.5 Pearson Correlations Between Raters’ Speech Ratings and Their Responses to the Social Attitudes Questionnaire.....	80
Table 4.1 Means (Standard Deviations) for Citizenship Labels by Speaker Group, Condition, and Location.....	110

Table 4.2 *Summary of Mixed-Effects Models for Citizenship Labels in the Nonreligious and Religious Conditions by Location and Speaker Group*.....111

Table 4.3 *Means (Standard Deviations) and Between-Context Comparisons for Raters' Responses to the Social Network Survey*.....114

Table 4.4 *Pearson Correlations Between Raters' Citizenship Evaluations and Responses to the Social Network Survey*.....115

Definition of Key Terms

Monolingual. For the purposes of this study, a monolingual English speaker is defined as someone who has used English as their primary language for most of their lives (including their education), and who continues to use almost exclusively English in their daily lives, including at home. This does not necessarily mean that English is the person's language from birth.

Bilingual. For the purposes of this study, a bilingual English–French speaker is someone who has spoken both English and French for most of their lives, and who continues to use both languages in their daily lives and at home. Proficiency in both languages is assumed, but this does not necessarily mean that the person has spoken both languages from birth.

CHAPTER 1

General Introduction

Throughout the years, the nation of Canada, as we are aware, became a land of immigration, a home to millions of people from different lands, ethnicities, cultures and religious beliefs. Every new comer that has settled in Canada in the last few hundred years and those who will settle in the future share one common name: “immigrants.”
(*Samer Mazjoub, Egyptian immigrant, President of Canadian Muslim Forum*)

Our attitudes control our lives. Attitudes are a secret power working twenty-four hours a day, for good or bad. It is of paramount importance that we know how to harness and control this great force. (*Irving Berlin, Russian immigrant to the United States*)

Canada welcomes about 300,000 immigrants per year (Government of Canada, 2022) and is globally perceived to be an open, multicultural society. It houses ethnically and linguistically diverse cities, such as Montreal, Calgary, Vancouver, and Toronto, and the nation as a whole has a solid history of effective approaches to immigrant policy that are regarded as holistic, neutral, thoughtful, as well as capable of linking commitment to action (Hiebert, 2016). Nevertheless, the country is not without intergroup tension. Indeed multiculturalism, though evidential of diversity and group recognition, has been connected to increased stereotyping and, subsequently, increased discriminatory behaviour, or biases between various ethnic groups (Kymlicka, 2009). Biases spring from beliefs about particular out-groups that are not rooted in fact, and they can develop about a particular individual or about a group of people as a whole. Since Canada is rapidly increasing its numbers of immigrants and its cultural diversity, it follows that discrimination, stereotyping (and therefore negative biases) may also be growing within its borders. According to the 2019 Race Relations in Canada survey (Neuman, 2019), 95% of Canadians believe that visible minorities in Canada experience racism, and 23% of Canadians report experiencing discrimination themselves, with racism being the most prevalent form

(Godley, 2018). Although attitudes differ across provinces, the results of the 2011 Canadian Election Study (Fournier et al., 2015) revealed that Whites were the most liked ethnic group, while Muslims were the least. While Quebecers have been documented to hold the belief that racism is decreasing (Neuman, 2019), they showed the most pronounced negative attitudes toward Muslims (followed by the Prairie provinces, including Alberta), a pattern that has been echoed in additional research (Angus Reid, 2017; CROP, 2017). This is but one example of how social attitudes toward varying ethnicities and religions might differ from province to province.

One way to explore these attitudes across multiple provincial contexts is through a comparison of Montreal and Calgary, two cities with comparable immigrant populations of 23% and 29%, respectively, but with different sociolinguistic contexts. Whereas 90% of Calgarians are English monolinguals, 57% of Montrealers are French–English bilinguals (Statistics Canada, 2017). Therefore, native-born residents in each context could be expected to view integration and acculturation from different perspectives. Nonetheless, the two cities also share the distinction of being more progressive than their respective provinces. This divergence is exemplified in leadership choices, for example. In 2010, Calgary elected Naheed Nenshi, who was the first Muslim to become a mayor in North America. More recently, Nenshi was replaced by Jyoti Gondek, the city’s first female mayor. Montreal also currently hosts its first female mayor, Valerie Plante, a longtime advocate for immigrants. In addition, the two cities have similarly prioritized social issues such as homelessness and overall community responsibility, despite drastically different forms of governance (Smith, 2018). These parallels suggest that despite differences in politics and problem-solving approaches, the two contexts are united in their social awareness of such matters as racial, ethnic, and linguistic discrimination.

These factors make Calgary and Montreal an interesting pairing for comparison and as East–West representatives of Canada as a whole. Through investigations in these two linguistic contexts, the overarching goal of this dissertation is to explore the perspectives of native-born residents—as representatives of their individual cities and of Canada more generally—about immigrants, most of whom are second language (L2) speakers in their communities. Of particular interest is how residents judge the speech of various L2 groups in terms of linguistic factors (comprehensibility, accentedness), status and solidarity traits (intelligence, friendliness, and trustworthiness), and citizenship status as well as how residents interact with L2 speakers and what their social attitudes are toward immigrants overall.

Social Bias and Language

Although much discrimination that exists can be attributed to appearance, accent is a salient marker of ethnicity (Hogg, 2006) leading to employment discrimination, stereotyping, and harassment (Munro, 2003). Often framed as a key component of cultural capital, accent accompanies features such as body language and humor in a “personality package” assessment valuable to prospective employers (Nghia et al., 2020, p. 42) and is, as Bourdieu (1977) notes, “an index of authority” (p. 653). Indeed, Pietraszewski and Schwartz (2014) assert that accented speakers are spontaneously and implicitly categorized by others in the same way that society categorizes people by gender or race. In their study, listeners spontaneously categorized speakers—independent from accent familiarity or ease of processing—based on their individual accent. However, just as society categorizes people by race, then assumes specific characteristics

to be true of all members of that race, a negative experience with one L2 speaker can lead to assumptions that extend to all speakers of that language.

However, not all societal groups categorize language similarly, which leads to subsequent differences in language attitudes. To illustrate this, Dragojevic et al. (2018) studied language attitudes toward varieties of Georgian (Tbilisi-accented Georgian and Mingrelian-accented Georgian) among listeners from three different ethnolinguistic groups which included Georgians, Armenians, and Azerbaijanis. In a matched-guise experiment, designed to extract listener reactions toward specific accents or dialects, the listeners rated samples produced by the same bilingual male speaker in terms of status (e.g., educated, cultured) and solidarity (e.g., attractive, amusing), then identified the speaker's Georgian origin (i.e., region or town). Georgians showed the highest accuracy in identification of the Tbilisi versus Mingrelian accent, which was tied to the more favourable or less favourable ratings overall for the Tbilisi and Mingrelian guises, respectively. In other words, not only do social groups differ in their attitudes toward the same language groups, but those differences in categorization are the driving force between evaluations of status and solidarity.

In Lev-Ari and Keysar's (2010) study, 30 native English-speaking listeners evaluated L2 speakers' credibility using recorded trivia statements. The researchers found a negative correlation between strength of accent and level of credibility, such that those with the heaviest accents were perceived to be the least credible. Furthermore, in a follow-up experiment, when told that "the experiment was about the effect of the difficulty of understanding speakers' speech on the likelihood that their statements would be believed" (p. 1095), listeners—despite some adjustments to their ratings for mildly accented speakers—still rated statements made by heavily accented speakers as the least credible. Although subsequent research has struggled to replicate

these findings in other linguistic contexts (Foucart et al., 2020; Stocker, 2017), there is ample evidence to support a connection between accent and judgments of status and solidarity in a variety of settings (e.g., Frumkin & Stone, 2020; Mirshahidi, 2016; Rakić et al., 2011).

In the context of daily life, judgments of such attributes as credibility and employability can be particularly harmful when assigned to L2 speakers. Nevertheless, these judgments are pervasive in society (Akehurst et al., 2018; Lippi-Green, 1994; Matsuda, 1991; Nguyen, 1993; Smith, 2005). For example, Timming (2017) investigated the effect of foreign accents on applicants' employability ratings and found that when listeners (223 US survey respondents with management experience) rated 20 L2 English samples of mock job interviews representing five different language groups (Indian, Chinese, Mexican, British, and American), they rated speakers with Indian, Chinese, and Mexican accents to be less employable and, if employed, less suitable for customer-facing jobs than speakers of American and British English. Similar results have been found in other work demonstrating that a foreign accent can negatively overshadow actual ability (Carlson & McHenry, 2006; Hosodo & Stone-Romero, 2010).

Though often linked to judgments of status and solidarity, accent categorization (and the social attitudes that stem from it) also surfaces in assessment contexts. In another matched-guise experiment, Ford (1984) showed that teachers in the southwestern United States evaluated the English writing samples of third and fourth graders (all native English speakers) more negatively when the work was paired with speech samples of Spanish-accented speakers than speech samples of native English speakers. Echoing this pattern, English lecturers at a Slovenian university self-reported biased assessment of student writing, citing knowledge of the student's L2 speaking proficiency as a key factor in their assessments (Sokolov, 2014).

Language-Based Bias in Canada

Canada, despite being a generally welcoming environment for newcomers, has seen its own struggle with negative attitudes toward L2 speakers, particularly regarding language and accent biases and the stereotyping and negative social attitudes that result from those judgments. In a Vancouver study of African migrants, for instance, Creese (2010) found that many had been repeatedly subject to accent discrimination. This treatment is thought to be manifested by what the author calls “an erasure of linguistic capital” (p. 300), where status acquired from being English-proficient in Africa was significantly reduced (due to the presence of an accent) in the new Canadian context. Prior to this, Dion (2001) found evidence of similar discriminatory attitudes experienced by Jamaican, Pole, and Somali immigrants in Toronto through the investigative *Housing New Canadians Project*. More recently, Baquiran and Nicoladis (2019) asked 161 University of Alberta undergraduates to evaluate recorded doctor–patient conversations in which the doctor spoke either Canadian- or Chinese-accented English. The Chinese-accented doctor was rated as significantly less competent than the native-speaking doctor, regardless of listener ethnicity and despite the fact that half of the listeners were Chinese Canadian themselves.

Though Canada, like many countries, has developed policies to protect immigrants from such discrimination, Oxman-Martinez et al. (2012) note that there is a disconnect between theory and practice. In their study of perceived peer and teacher discrimination among 1,053 adolescents in six major Canadian cities, the researchers found that more than a quarter of the 11- to 13-year-olds surveyed felt that they had been subject to hostile, discriminatory behaviour in the month prior to the study by peers or teachers based on their place of birth, physical

appearance, or language. Furthermore, a sense of perceived discrimination led to lower self-esteem and feelings of lower social and academic competence.

Certainly, the two major Canadian cities targeted in the current study, Calgary and Montreal, have experienced their share of biased attitudes toward immigrants and languages. In a series of ethnographic interviews conducted with 11 female immigrants who had recently settled in Calgary, Graham and Thurston (2005) found that participants attributed their employment challenges to immigrant status rather than qualifications, and that the challenges of being an immigrant and a L2 speaker extend to various facets of everyday life in Calgary. For example, Oppong's (2018) study of Ghanaian immigrants found that many perceived their accent to function as a barrier to effective communication with physicians, leaving many patients to avoid the healthcare system altogether.

Social bias against linguistic groups, however, can even exist in majority-language environments. Exploring the language attitudes of four mother-tongue language groups—French, English, bilingual French–English, or other—Kircher (2012) asked 160 Montreal CEGEP students, 63.1% of whom were either immigrants or descended from immigrants, to complete a questionnaire regarding general attitudes toward Quebec French and European French, then rate 15 matched-guise speech samples (five each Canadian English, Quebec French, and European French) for characteristics of status (e.g., intelligence, leadership) and solidarity (e.g., kindness, likeability). Results of the questionnaire and the speech ratings revealed more favourable attitudes toward European French than Quebec French regarding status, despite a sizable proportion of the participants being Quebec natives.

Biases experienced by immigrants in Canada are not limited to language, however. In a study involving thousands of mock resumes sent to various companies in three major Canadian

cities, Dechief and Oreopoulos (2012) found that in Montreal, applicants with the appropriate education and expertise—even when acquired in Canada—were 40% more likely to be called for an interview if the name on the resume sounded English (e.g., Matthew or Philip) than if the name sounded foreign (e.g., Samir or Wang), and the effect worsened if a foreign name was also paired with foreign education and expertise. This bias (conscious or unconscious) often stems from issues such as lack of training regarding hiring practices and a perceived (but unjustified) lack of fit for the company based on stereotypes about the skills of certain ethnic groups (Esses et al., 2006). Often compounding these negative biases are extralinguistic markers of ethnicity, such as religion. Vang et al. (2019), for instance, found that as many as 24.8% of Canadian immigrants who practice minority religions (e.g., Sikhism, Islam, Judaism) had been discriminated against based on their religion in the five years preceding the study, and many equated the effects of such discrimination to other significant life trauma (such as losing employment or a spouse). In sum, what is clear from prior research is the need for continued study and education toward reducing biases that interfere with equality.

This call-to-action springs not only from empirical research, but also from recent media accounts of ongoing discriminatory acts against people of ethnic, religious, and linguistic out-groups carried out in real-life contexts across Canada, including the two wider contexts targeted in the current study. In Calgary, for instance, a Muslim mother and her children were verbally accosted while walking downtown (Gilligan, 2021); another Calgary woman was chased by a man shouting racial slurs when she refused to hand over her coffee (White, 2021), and racist graffiti defaced the walls of the Sikh Society of Calgary (Bartko et al., 2021). Meanwhile, in Montreal, where racially-motivated hate crimes have increased by 53% since 2020 (Lofaro & Lurie, 2021), former mayoral candidate Balarama Holness reported receiving racist and

threatening messages throughout his campaign (Lalonde, 2021), and the rise in anti-Asian violence has had immigrants considering relocation (Brownstein, 2021). Not only do these media accounts reveal the urgency for bias reduction, but the headlines themselves may increase intergroup anxiety, thereby negatively influencing attitudes toward immigrants (Atwell Seate & Mastro, 2015).

The Dissertation Studies

Set against this backdrop, this dissertation aims to explore how native-born Canadian residents judge L2 speech across a variety of dimensions including global speech measures (comprehensibility, accentedness), status and solidarity traits (intelligence, friendliness, trustworthiness), and belonging (citizenship labeling). Of particular interest is how sociolinguistic context (predominantly monolingual vs. bilingual) might influence ratings and whether rater beliefs, namely social attitudes toward immigrants and social network exposure, are associated with ratings of L2 speech in situations where it is experienced with and without visual information revealing a speaker's ethnic origin and religious affiliation. **Study 1** investigates native-born residents' judgments of speaker comprehensibility and accentedness when speech samples are presented as audio-only, audio paired with a matching ethnic image, and audio paired with a mismatching image. While investigations of L2 comprehensibility and accentedness as judged by naïve raters are plentiful (e.g., Saito et al., 2016; Taylor Reid et al., 2019), and many studies employ a matched-guise technique (e.g., Hansen et al., 2017b; Rubin & Smith, 1990), there are three major distinguishing aspects of Study 1. First, ratings are compared in two different sociolinguistic contexts (predominantly monolingual Calgary and predominantly bilingual Montreal), a comparison designed to further our understanding of how rater background and environment can affect judgments of L2 speech. Furthermore, whereas many

matched-guise studies focus on one or two L2 speech varieties, Study 1 includes four (Arabic, Mandarin, Punjabi, Tagalog), with North American English as a baseline, which allows for a more representative sample of what native-born Canadians encounter in daily life, at least in major cities. Finally, there is also limited prior research that explores whether raters' social attitudes toward immigrants are associated with their judgments of L2 speech, and certainly none that examines this relationship across sociolinguistic contexts in Canada.

While Study 1 focuses on global measures of L2 speech, **Study 2** focuses on judgments of status and solidarity traits (intelligence, friendliness, trustworthiness) among the same raters and using the same L2 speech samples. This time, however, speech is presented alone or with either a matching ethnic image or a matching religious image (i.e., a speaker wearing a religious head garment such as hijab or turban). Although there is a small body of research that has incorporated religion into matched-guise experiments (e.g., Lybaert et al., 2022), this design is rare in Canadian contexts and does not exist at all with the same range of L2 speaker group representation. Religion is a particularly important aspect to consider in Canada, where religious discrimination is rising and the nation heatedly debates the justness of Quebec's Bill 21, which prohibits people in public service positions from wearing religious symbols while fulfilling their duties. Study 2 also pools ratings from Montreal and Calgary to provide one robust, multi-city sample with East–West representatives of Canada's two primary sociolinguistic environments (monolingual and bilingual). In doing so, this is one of only a handful of studies that examines judgments of L2 speech in multiple Canadian contexts. Finally, Study 2 explores associations between social attitudes and ratings of L2 speaker status and solidarity, an investigation that is limited in prior research, particularly as it pertains to understanding rater attitudes toward

acceptance of immigrants in prominent societal roles as well as perspectives about immigrant civic and political involvement, all of which factor heavily into Canadian national identity.

One aspect of human behaviour that has been shown to influence how people perceive and treat others is social interaction, which is explored in relation to judgments of L2 speech in **Study 3**. For example, the more contact a person has with people of diverse cultures, the less likely it is for that person to develop negative biases. This idea springs from a general understanding in sociology and psychology that regular interaction with people of different languages, ethnicities, and faiths, for instance, leads to greater acceptance of diversity in general (Hewstone, 2015; Tropp & Pettigrew, 2005). Social networks can, therefore, be a valuable tool in investigating the roots of individual and group linguistic biases, particularly in contexts that lie at the far ends of the continuum, where contact with L2 speakers is either restricted or robust. In other words, investigating social networks alongside ratings of L2 speech can sharpen our understanding of how listener exposure shapes the way others are perceived in terms of their language abilities, status and solidarity, and overall belonging. Therefore, one goal of Study 3 is to compare the L2 speaker social networks of raters in predominantly monolingual Calgary and predominantly bilingual Montreal. This is the first study to compare networks in two Canadian contexts and explore the relationship between network characteristics and ratings of L2 speech.

Another largely unexplored but vital aspect of L2 speaker acceptance in immigrant-hosting environments is how native-born residents perceive minority group members' belonging, particularly in terms of citizenship. Study 3 aims to address this gap. Although many L2 speakers feel they belong in Canadian society, their language status may unfairly limit their involvement (Boyd, 2009). Given that belonging is such a vital component of immigrant success in a host community (Benish-Weisman & Horenczyk, 2010; Berry & Hou, 2017), Study 3 explores how

native-born residents perceive the citizenship status of speakers from four L2 speaker groups, when their speech is presented through audio alone or with either a matching ethnic image or a matching religious image. This is the first known study to investigate native-born residents' judgments of L2 speaker citizenship in North America.

In sum, the existence of negative social biases—formed when beliefs about out-group members are not rooted in fact—is well-documented, and prior research has offered rigorous evidence of how those biases might be formed. What remains unclear is a thorough understanding of the effects of contextual factors—such as linguistic environment and exposure—on social attitudes toward L2 speech, a sharpened understanding of how we might best explain this phenomenon from a theoretical perspective, and the relative impact of factors such as speech, ethnicity, and religion on attitudes toward immigrants. This dissertation explores the specific roles of social networks, exposure, and context in the formulation of biases as well as both the relationship between these factors and their effect on native-born residents' ratings of immigrants' speech. It also aims to capture a multi-city Canadian perspective of relationships between native-born residents and immigrants. Unique to this endeavor is the exploration of how native-born residents view L2 speakers in terms of belonging—whether their presence is temporary or permanent, and whether they are entitled to the same rights as Canadian-born citizens. Armed with this understanding, future recommendations can be made as to how negative biases might best be mitigated through such strategies as intergroup contact and, more specifically, communicative accommodation—a key component in establishing common ground and equitable interlocutor power.

CHAPTER 2

Study 1: The Role of Sociolinguistic Context in Ratings of Immigrants' Second Language

Speech: A Look at Calgary and Montreal

The world exists in a constant state of change as individuals move from place to place either by choice, as with skilled immigrants who relocate for better employment opportunities, or by necessity, as with those seeking asylum. Within North America, a continent with immigrant numbers that are climbing faster than any other (Esses & Abelson, 2017), Canada exists as a welcoming society for those seeking opportunity or refuge. To accommodate increasing numbers of immigrants, however, Canada has had to reimagine how best to provide support to new arrivals, benefit from an influx of new employment skills and cultural ideas, and create a harmonious environment for both the country's citizens and its newcomers (Joshee & Thomas, 2017). In addition, although national policies—such as the 2002 Immigration and Refugee Protection Act—exist to manage an ever-changing ethnic landscape, provinces have adopted laws of their own, which may or may not be in alignment with each other. Considering that the majority of Canada's foreign-born population is distributed across metropolitan areas in British Columbia, Ontario, Alberta, and Quebec (Romaniuk, 2017)—provinces with, as Blake (2019) notes, four markedly different provincial identities—it would be difficult to create a one-size-fits-all immigration policy. However, even to establish suitable policy at the provincial level, there must first be a clear understanding of attitudes native-born citizens hold toward immigrants.

The Roots of Group Conflict

In the same way that Tajfel (1972) noted that groups compete for power and status as part of their social identity, Stephan and Stephan (2000) assert that they might also be in competition

for cultural resources, which broadly refer to those things that a social group has deemed to be of value (e.g., jobs, wealth), and that perceived threat to those resources is a strong predictor of group attitudes that can exacerbate discrimination toward an out-group (Albarello & Rubini, 2018). Some conflicts arise from the concrete existence of conflicting goals or competition which, according to realistic group conflict theory (Sherif, 1966), leads to intergroup hostility. One example of this might be distinct political advantages that one language group holds over another, as is the case in Quebec, where Francophones possess greater political power than Anglophones. But threats are not solely based on actual disparity; in fact, perception of inequality can cause even greater discord (Kimmelmeier & Winter, 2000; Rouhana & Fiske, 1995). According to the instrumental model of group conflict (Esses et al., 1998), when resources are believed to be limited, scarce, or unevenly distributed, a threatened group may seek to eliminate the source of competition. Presumed threat may thus result in negative attitudes toward immigrants, as even high-status groups might feel threatened if they sense that a possible shift in hierarchy is imminent.

At this moment in human history, as common resources are rapidly depleting, tensions between competing groups are particularly high (Hodson et al., 2013). For instance, one group might try to decrease perceived value of a competitor by highlighting that group's negative attributes or attempt to increase the perceived value (and therefore competitiveness) of their own group by highlighting its positive attributes. In essence, people with a strong sense of group identity might emphasize the differences between their own group and others. Another, more extreme strategy is for a group to decrease proximity or deny access to a competitor, such as when a government denies immigration. Sioufi and colleagues (2015) connect these perceptions and behaviours to the existence of "zero-sum" beliefs, a game theory first introduced by Von

Neumann and Morgenstern (1944), which posits that when resources are assumed to be limited, an individual or group gain of any resource results in the loss of that resource from another. In one example of how zero-sum beliefs can lead to negative attitudes toward immigrants, Wilson (2001) analyzed 1,474 responses of US residents in three geographical areas related to perceived consequences of immigration and found that nearly all Americans believed immigration to be threatening on multiple levels: realistic (e.g., job-related), symbolic (e.g., related to national unity), and zero-sum (e.g., related to the distribution of wealth). Louis et al. (2013) found similar results among English-speaking citizens of Australia and Canada who held negative and distrustful attitudes towards immigrants, believing themselves to be in direct competition with linguistic minorities for limited resources, both objective (e.g., jobs, language services) and symbolic (e.g., values, religion).

Exploring Ethnic and Linguistic Biases

One challenge in the investigation of biases toward immigrant speakers is to disentangle the various factors that trigger negative reactions. For instance, if negative attitudes are rooted in fear that the arrival of immigrants will alter the way in which business is conducted, the intrusion of foreign languages might be of greatest concern. If there is a fear that the values of a community might be negatively influenced, immigrants' diversity in cultural practice might be more disturbing. Given that an individual's social identity can derive from multiple groups simultaneously (e.g., young Indian female), it is difficult to isolate which group identity might fuel a negative response from others; however, a small body of prior research has attempted to investigate the way identities of various minority groups intersect and overlap. For example, Nadal et al. (2015) examined intersectional microaggressions (across the domains of race, ethnicity, gender, sexuality, and religion) using qualitative secondary analysis to revisit data

from pre-existing discrimination research. They found that across all studies, participants experienced microaggressions based on both their singular identities (e.g., Filipino, lesbian) and their intersectionalities (e.g., Black Muslim, Asian female). To complicate matters of disentanglement and mitigation, these negative attitudes are often unconsciously held (Nadal et al., 2014; Walls et al., 2015).

One way in which ethnic and linguistic attitudinal biases have been examined is through ratings of second language (L2) speech which (among various dimensions) have targeted the constructs of comprehensibility, or listeners' perceived difficulty in understanding L2 speech (Munro & Derwing, 1995), and accentedness, or how much L2 speech differs from the variety spoken by native speakers (Munro & Derwing, 1995). Prior research has established comprehensibility as the key factor in successful interaction between those who speak English as a primary language and those who do not (Derwing & Munro, 2009, 2015), and accentedness has been widely acknowledged as a salient marker of linguistic otherness (Kang et al., 2010). In one example of how speech ratings might indicate biases, Hansen and Dovidio (2016) had 65 US adults evaluate Mandarin- and Spanish-accented speech samples (identical in content) for speaker comprehensibility and hireability while also measuring these participants' views of their own ethnic group's social dominance. The participants' pre-existing biases appeared to contribute to perceived difficulty of understanding L2 speakers, leading to overall negative attitudes and lower hireability ratings, where biases and comprehensibility likely reinforced each other in a cycle of discriminatory factors that is difficult to break. In another example of how speech ratings relate to social biases, De Souza et al. (2016) asked 71 Portuguese university students to evaluate identical mock job interview excerpts from either a Portuguese- or a Brazilian-accented male speaker, then evaluate the speaker's hireability. The Brazilian-accented

speaker received significantly lower hireability ratings, supporting the assumption that negative attitudes toward Brazilians as an immigrant group in Portugal would extend to judgments of their employability, based on their accent. The important message emerging from these findings is that both social perceptions about L2 speaker groups and the actual quality of an individual's L2 speech can contribute to how immigrants are perceived in society.

The connection between pronunciation factors and biases can, of course, extend beyond employment contexts. One way in which researchers have investigated this relationship is through matched-guise technique, first introduced by Lambert et al. (1960), to elicit cultural stereotypes (perhaps either suppressed or subconsciously held) about speakers whose speech is marked by a nonnative or nonstandard accent. Though initially introduced as an audio-only match/mismatch experiment, more recent use of photo guise has allowed researchers to examine not only attitudinal responses to varying degrees of nonnativeness, but also to activate any pre-existing group stereotypes based on ethnicity without having to explicitly ask participants about their attitudes. Results can reveal important information about how in- or out-group membership is determined based on whether listeners are rating the speech itself or whether they are influenced by visual information about a speaker's ethnicity.

In one early photo guise experiment, as a way of exploring student biases against L2-speaking teaching assistants, Rubin and Smith (1990) asked 92 university students who were born in (and had never left) the US to listen to speech samples of various levels of accentedness. The samples, which were recorded by a supposed university instructor discussing one of two academic topics (humanities or science), were paired with a photograph of either a Euro-American (White) woman or an Asian woman. Not only did students judge more accented speakers to be poorer teaching assistants but, when humanities-themed speech samples were

paired with an Asian face, students also reported lower comprehension of the material. This suggests that listeners' judgments of L2 speakers may reflect not only the speakers' linguistic performance but may also stem from beliefs about particular ethnic groups, as triggered by images. In another example, Kang and Rubin's (2009) study of 158 culturally diverse individuals (i.e., native and L2 speakers with various levels of exposure to, education in, and experience with English) found that listeners with less L2 exposure in their daily lives rated the same English speech sample less favourably in terms of quality of instruction and perceived "standardness" of accent and understood less of what was being said when the sample was paired with an East Asian face rather than a White face. In this case, results were affected by two factors beyond the speech itself: visual information and the listeners' own linguistic exposure.

While listeners' linguistic exposure may include various travel, study, and socialization experiences, the place where they make their home—for instance, in a predominantly monolingual or predominantly multilingual environment—is a key determinant in shaping their own linguistic identity (Green & Abutalebi, 2013). In monolingual contexts, speakers of a majority language remain the dominant population, even when there is a healthy influx of immigrant languages. Other contexts, however, illustrate what Siegel (2003) calls a co-existing L2 environment in which more than one language enjoys majority status. The monolingual versus multilingual status of the host environment can be an important factor in how individuals respond to other ethnolinguistic groups in that it helps them to create an acceptance of their own multicultural identity that translates to a reduction in prejudice against others (Guboglo, 1974). In one example of this, Singh et al. (2020) found that bilingual preschoolers showed less bias against individuals of African descent (presented through images) than their monolingual classmates. Bilingualism has also been shown to foster acceptance of other ethnic groups

because it raises one's cultural awareness and inevitably leads to positive intergroup contact (Genesee & Gándara, 1999). In fact, the very belief that one can successfully become bilingual and multilingual, which Lou and Noels (2020) refer to as a “growth language mindset” (p. 51), can lead to greater tolerance of newcomers and increased acceptance of cultural diversity, especially since linguistic assimilation can be a primary factor in acceptance by a host community (Adserà & Pytliková, 2016).

The Current Study

In sum, biases against L2 speakers can be captured through ratings of such linguistic factors as comprehensibility and accentedness, where judgments about the same speaker may be harsher or more lenient based on a listener's exposure to, fear of, or pre-existing beliefs about, certain immigrant groups as activated, for example, by ethnic images that accompany the actual speech heard (e.g., Kang & Rubin, 2009; Rubin & Smith, 1990). However, it is also important to consider how biases toward immigrants are shaped by linguistic context (i.e., monolingual vs. multilingual), especially because—aside from a small body of work focused on local issues or international comparisons (e.g., Berry, 2006; Esses et al., 2006; Wilkins-Laflamme, 2018)—there have been no cross-context comparisons between different Canadian sociolinguistic environments. Therefore, this study's overall goal was to examine whether native-born residents of two distinct Canadian cities (Calgary vs. Montreal) differ in their evaluations of immigrants' speech (in terms of its comprehensibility and accentedness) and in their views of immigrant roles in society as a function of their sociolinguistic environment (predominantly monolingual vs. multilingual).

In the sociolinguistic landscape of Canada, Calgary and Montreal are suitable representatives of the clear contextual distinction between predominantly monolingual and

multilingual host environments for immigrants. Although the two cities are comparable in terms of overall population (Montreal 1.8 million, Calgary 1.3 million) and immigrant population (Calgary 29%, Montreal 23%), 70% of Quebecers who claim English as their primary language (the majority of whom reside in Montreal) are bilingual English–French (Department of Canadian Heritage, 2012), whereas only about 7% of English-dominant Calgarians regularly interact in both English and another language (Statistics Canada, 2017). Despite these differences in language use, cultural diversity is strong in both contexts. Montreal, for instance, is represented by 120 different cultural communities (Immigration, Francisation et Intégration Québec, 2019), and more than 85% of Quebec’s immigrants have chosen Montreal as their home (Palardy & Filip, 2015), which accounts for Montreal’s description as a “highly heterogeneous, multi-racial and multi-ethnic constituency” (Department of Canadian Heritage, 2012, p. 8). The primarily monolingual city of Calgary has an even more diverse population, however, as it is represented by over 240 different ethnic backgrounds (Statistics Canada, 2017). Thus, even though residents in both cities have the opportunity to engage in frequent multicultural interactions and to acquire large, culturally diverse social networks, one key difference between the two cities is their predominantly multilingual (Montreal) versus monolingual (Calgary) setting.

To carry out a Calgary–Montreal comparison of attitudes toward immigrants, a total of 340 native-born residents of Calgary and Montreal were presented with speech samples recorded by recent immigrants from four different linguistic backgrounds, representing either the top or fastest growing immigrant languages in common between the two cities (Arabic, Tagalog, Mandarin, and Punjabi), with North American English used for comparison. These residents evaluated the audio recordings for comprehensibility and accentedness (among other dimensions

not focused here), where sometimes the target speech samples were presented alone and sometimes they were accompanied by images that either matched or mismatched the speaker's ethnicity, as in typical matched photo guise experiments (e.g., Kang & Yaw, 2021; McGowan, 2015), so that speech ratings could be explored in situations where the image prompted potential biases that were congruent or incongruent with the speaker's identity, relative to the no-image condition. The same native-born residents also completed an extensive social attitudes questionnaire targeting their attitudes towards immigration and the role of immigrants in society, capturing their overt opinions about immigration. Of key interest was the comparison between the Calgary and Montreal raters' speech ratings and their responses to the social attitudes questionnaire, as well as a potential relationship between these measures. The following research questions were explored:

1. How do ratings of L2 speakers' comprehensibility and accentedness compare for native-born residents of Montreal and Calgary when speech samples are presented in three conditions (audio-only, audio paired with a matching ethnic image, and audio paired with a mismatching ethnic image)?
2. How do native-born residents of Montreal and Calgary compare in how they view immigrant roles in society and the government's response to immigration, as revealed through questionnaire responses?
3. Is there a relationship between Montreal and Calgary residents' attitudes toward immigrants (shared in their questionnaire responses) and their ratings of immigrants' comprehensibility and accentedness?

Based on the Calgary–Montreal contextual differences in linguistic environment, for the first research question, it was expected that residents of Calgary and Montreal would differ in

their ratings of L2 speech, such that Montrealers would assign more favourable ratings than Calgarians, evaluating immigrants as less accented and more comprehensible, and that at least some of those differences might be attributable to Montreal residents' bilingualism and their pre-existing views about particular ethnolinguistic groups (as revealed through their responses to the social attitudes questionnaire). Montrealers, many of whom are speakers of two majority languages, who possess a multicultural identity, and who have already demonstrated high value in the workforce (Da Silva et al., 2007; Levine, 1988), might have less to fear from the arrival of immigrants than Calgarians, because Montrealers have experienced first-hand the benefits of linguistic diversity. It was also expected that those whose languages are more familiar to the residents in Calgary (Tagalog, Punjabi, and Mandarin) and Montreal (Arabic), based on the relative weight of the groups in the ethnolinguistic landscape of each city (Statistics Canada, 2017), would be judged as less accented and more comprehensible than the remaining speaker groups. Finally, if the ratings differed for the same speech sample when presented as audio-only versus with a matching/mismatching ethnic image (e.g., an Arabic speaker paired with a Middle Eastern vs. a White face), then it would be clear that suggestions of ethnicity were activating a response to that particular ethnic group beyond judgments of the speech alone. These differences were expected to be magnified in comparisons of the audio-only versus match conditions, compared to the audio-only versus mismatch conditions, because the matching cues available through speakers' speech and their visuals might reinforce each other, with raters reacting to speakers based on the ethnicity presumed from both sources. Due to the lack of prior cross-context comparisons, we refrained from making refined predictions as to potential differences between Calgary and Montreal residents' responses to immigrant speech paired with a matching or mismatching ethnic image, relative to the audio-only condition.

For the second research question, based on social identity theory (Tajfel, 1972), instrumental model of group conflict (Esses et al., 1998), and zero-sum theory (Von Neumann & Morgenstern, 1944), it was expected that the residents in both contexts would exhibit some level of overt negative attitudes toward immigrants. Although many Canadians believe immigration to be good for the country, the perceived threat of newcomers to available resources persists (Soroka & Robertson, 2010). It was further expected that residents who spend the most time in a predominantly monolingual environment such as Calgary would exhibit more pronounced negative attitudes toward immigrants, and that many residents of Montreal would be more generous in their social attitudes, not only due to their own dual-language identity, but also due to the exposure they experience in a multilingual environment. Finally, for the third research question, it was expected that any biases against particular immigrant groups expressed through the social attitudes questionnaire would be associated with harsher comprehensibility and accentedness ratings of L2 speakers, with the harshest ratings coming from those who hold the strongest negative attitudes.

Method

Raters

Raters included 340 native-born residents (170 Calgarians, 170 Montrealers) between the ages of 20 and 60, evenly distributed across decades. In all, there were 200 females, 131 males, and nine non-binary individuals. All used English as the primary language of interaction in their daily lives, though not all were native speakers of English (15 L2 speakers and 155 native speakers in each context). The choice to seek raters for whom English is the predominant language of communication, particularly in the multilingual setting of Montreal, was made to facilitate a clear comparison between contexts and to sidestep any English proficiency issues for

L2-dominant raters, resulting in rater exclusion due to insufficient aural skills to judge L2 English speech. In addition, raters were born in the target environment and had been living there for at least three years prior to the study. Therefore, the rater group included lifetime residents as well as those who had spent periods away from the target context but returned, with an average length of residence of 25 years for Montrealers and 32 years for Calgary. Raters were recruited through social media, emails to local community organizations, word of mouth, and participant referrals. In both contexts, raters were similar in their daily interaction with native English speakers, as reported on a 0–100 percentage scale ($M_{\text{Montreal}} = 86.00$, $SD = 17.99$; $M_{\text{Calgary}} = 82.93$, $SD = 16.86$). However, raters differed in the number of languages spoken, such that Montreal raters reported speaking more languages than Calgary raters ($M_{\text{Montreal}} = 2.54$, $SD = 0.91$; $M_{\text{Calgary}} = 1.58$, $SD = 0.90$), $t(338) = 102.61$, $p < .001$, $d = 1.06$. They also differed in their L2 proficiency for those who identified as multilingual. Using a 100-point scale, where 0 was labeled “can say a few words” and 100 was labeled “completely fluent,” Montreal raters estimated their L2 proficiency higher than Calgary raters ($M_{\text{Montreal}} = 84.60$, $SD = 14.56$; $M_{\text{Calgary}} = 61.27$, $SD = 33.78$), $t(338) = 77.28$, $p < .001$, $d = 1.05$.

Materials

Speech Samples and Photo Guises

To collect speech samples suitable for the study, 10 Montrealers—one male and one female from each of the five target language backgrounds (Arabic, Mandarin, Punjabi, Tagalog, and North American English), took part in an online mock job interview in English with an experienced interviewer. The L2 speakers, all born outside Canada, represented typical speech for the three most prominent and/or fastest growing immigrant language backgrounds in the two target contexts; native speakers of North American English provided a baseline comparison. All

L2 speakers were noticeably accented, and they self-reported intermediate to advanced speaking proficiency. From each interview, a 30–45 second segment in which speakers discussed how they respond to constructive criticism was extracted. In exchange for their contribution, participants received both compensation and feedback on their interviewing skills.

Photos paired with the speech samples were chosen from a pre-rating task that included images from the Chicago Face Database (Ma et al., 2015), the Bogazici Face Database (Saribay et al., 2018), the London Face Database (DeBruine & Jones, 2017), and from images of several volunteers recruited by the researcher. All images showed people with minimal augmentations to appearance (e.g., plain hairstyles, no tattoos, minimal make-up), and individuals' ages ranged from 24 to 32 ($M = 27.08$, $SD = 2.40$). To create an even more unified set, a Photoshop professional altered each image to show each person against the same neutral background and in the same grey t-shirt (see Appendix A for examples). Thirty-eight images (at least 3 male and 3 female of each target ethnicity) were presented to 10 Montrealers between the ages of 24 and 55 ($M = 36.80$, $SD = 11.43$), who were asked to identify the person's ethnicity from among five broad groups (see Appendix B for the full list), then rate each photo for prototypicality and traits of status (intelligence) and solidarity (friendliness, trustworthiness). The final set of 10 images (1 male and 1 female for each of the five target ethnicities) was determined based on the highest rated ethnic prototypicality for a male and female within each target ethnic group, with a score range of 65–84%.

Speech Ratings

Raters experienced the audio narratives in three conditions (audio-only, audio with a matching ethnic image, audio with a mismatching ethnic image), where each language group (Arabic, Mandarin, Punjabi, Tagalog, North American English) was represented by one male and

one female speaker. In the image conditions, each speech sample was paired with a photograph suggesting a potential visual depiction of the speaker, though none of the photos were of the actual speakers. To ensure that all possible pairings would be presented with the same frequency, 16 versions of the survey were created, and 10 raters were randomly assigned to each version per location (Calgary, Montreal), for a total of 320 raters. Across all survey versions, the speech samples from each language group occurred eight times with a matching ethnic image and eight times with a mismatching image from one of the four remaining ethnic groups for a total of 80 target items (5 language groups \times 16 voice–image pairings). In addition, 80 other voice–image pairings included male and female speakers presented with matching or mismatching ethnic images containing a religious symbol (e.g., hijab, turban), but these religious image stimuli served as fillers and were not used for analysis. However, because all possible combinations of voice–image pairings could not be presented in each survey version (so that a given rater would not be exposed to the same speaker’s voice more than once), the 320 raters (160 per location) were presented with a random subset of five image–voice pairings, contributing 1,600 datapoints to the dataset. To obtain baseline ratings, an additional 20 raters (10 per location) completed an audio-only version of the survey, contributing an additional 200 datapoints. In this version, raters heard all 10 speech samples (1 male and 1 female per speaker group) presented in randomized order, but this time, the rating instructions and scales were presented without an accompanying image. Across all conditions, no rater heard the same audio or saw the same image more than once.

Raters evaluated each speech sample for comprehensibility (ease of understanding) and accentedness (how much L2 speech differs from the variety spoken by native speakers) using a 100-point scale (for validation of similar scales, see Saito et al., 2017). For comprehensibility,

the scale was labeled at the left endpoint (corresponding to 0) as “very difficult to understand” and at the right endpoint (corresponding to 100) as “very easy to understand.” For accentedness, the scale was labeled at the left endpoint as “no accent” and at the right endpoint as “very heavy accent.” After hearing the speech sample one time (with no replay permitted), raters were asked to evaluate the speaker for comprehensibility and accentedness (both addressed in this study) as well as for friendliness, trustworthiness, and intelligence (discussed in a separate study). All scales began at the central mark (50) and, as raters clicked and dragged the toggle, numbers corresponding to each individual datapoint appeared above the scale. Some movement of each scale (even to return to 50) was required before raters could proceed to the next sample.

Social Attitudes Questionnaire

Adapted from the European Social Survey (2015), the social attitudes questionnaire (Appendix C) explored raters’ attitudes toward immigrants in terms of such topics as desired individual characteristics (e.g., “How important should it be for someone born, raised, or living outside of Canada to have work skills that Canada needs?”), desired race and religion (e.g., “How important should it be for someone born, raised, or living outside of Canada to be White?”), and the willingness for immigrants to hold prominent roles in everyday life (e.g., “How much would you mind or not mind if someone from another country who is a different race or ethnic group from most Canadian people taught your children in school?”). Participants rated the importance of each aspect using a scale from 0 to 100, with 100 representing either “very important” for the questions about individual characteristics or “not mind at all” for questions about immigrant roles in society. Raters also answered broader questions (discussed in a separate study) about their own views of Canadian culture, government policies, and the civic and political impact of immigrants. Similar to the comprehensibility and accentedness scales, all

scales began at the central mark (50) and, as raters clicked and dragged the toggle, numbers corresponding to each individual datapoint appeared above the scale. Some movement of each scale (even to return to 50) was required before raters could proceed to the next sample.

Procedure

Raters completed all tasks through LimeSurvey in a session that lasted approximately 30 minutes and included a consent form, a background questionnaire, a social network survey (the data from which are discussed in another study), the speech rating task, and the social attitudes questionnaire. The choice to use strictly online tools was born part out of practicality and part out of necessity due to COVID-19, which limited opportunities for in-person interaction at the time of data collection. Although online testing limits the researchers' control over some aspects of data collection including timing and noise, online tools have shown high internal consistency comparable to lab-collected data (Gosling et al., 2004; Nagle, 2019). For added security, however, several additional controls were added. Raters were not allowed to return to completed sections, change their answers, or skip audios or scales; their progress was time tracked, they were asked to use personal headsets for listening or complete the task in a quiet location, and they were encouraged to report any noise or interruptions that occurred during the rating session.

Data Analysis

Scalar ratings of accentedness and comprehensibility (0–100 scale) were first compiled into a matrix containing a total of 1,800 observations. As shown through two-way, consistency, average-measure intraclass correlations, raters demonstrated high consistency in evaluating comprehensibility (Calgary: .86–.94; Montreal: .83–.93) and accentedness (Calgary: .95–.99; Montreal: .96–.99) across the three rating conditions. Questionnaire responses (0–100 scale) were first checked to ensure that all items were consistent in their directionality, and responses to

two items were reverse-coded to achieve that. Based on both topic similarity and the patterning of responses in preliminary analyses, as confirmed by reliability indexes (Cronbach's α), similar questions (2–5 per topic) were grouped by theme and a mean value was calculated per rater across all questions within each theme, to produce a more robust set of responses per topic. For this study, three composite measures identified as most relevant to raters' general acceptance of immigrants as individuals were selected for analysis, with the remaining data reserved for future studies. The first measure included five questions about desired individual characteristics of immigrants, such as importance of education level, work skills, commitment to the way of life in Canada, and the ability to speak English or French ($\alpha_{\text{Montreal}} = .83$, $\alpha_{\text{Calgary}} = .80$). The second measure included two questions about the importance of being Christian or White ($\alpha_{\text{Montreal}} = .79$, $\alpha_{\text{Calgary}} = .83$), and the third measure included four questions about raters' willingness to accept an immigrant as their boss, physician, in-law, or child's teacher ($\alpha_{\text{Montreal}} = .87$, $\alpha_{\text{Calgary}} = .85$). These three derived measures of the raters' views of immigration were then incorporated in the data matrix.

To address the first and the second research questions, data analyses were carried out through mixed-effects modeling using the lme4 package (Bates et al., 2014) in R version 4.0.2 (R Core Team, 2020). Although there is no consensus regarding the criteria for considering statistical significance in mixed-effects modeling, with some scholars using $t > |2.00|$ as the benchmark of significance (Linck & Cunnings, 2015), we examined 95% confidence intervals (CIs) to check the statistical significance of each parameter (interval does not cross zero). To account for random variance, all models included random intercepts for participants (i.e., raters) and, where applicable, for items (i.e., speakers). Random slope models were also examined; however, the inclusion of random slopes did not improve model fit for any outcome variable, so

the final models excluded random slopes. To address the third research question, Pearson correlations were carried out to examine potential relationships between raters' views about immigrants and their ratings of comprehensibility and accentedness.

Results

Speech Ratings

The first research question asked if there are any differences between Calgary and Montreal raters' evaluations of comprehensibility and accentedness for L2 speech presented in three visual conditions. The ratings of comprehensibility and accentedness (summarized in Table 1) were analyzed separately for each speaker group. In each analysis, condition (audio-only, matching ethnic image, mismatching ethnic image) and location (Calgary, Montreal) served as fixed effects, and an intercept for rater and an intercept for speaker were included as random effects.

Table 1

Means (Standard Deviations) for Comprehensibility and Accentedness Ratings by Speaker Group, Condition, and Location (0–100 Scale)

Speaker	Calgary			Montreal		
	Audio-only	Match	Mismatch	Audio-only	Match	Mismatch
Comprehensibility						
Arabic	65.25 (25.14)	61.31 (26.97)	62.71 (23.94)	60.50 (27.68)	69.45 (24.25)	67.14 (25.07)
Mandarin	72.75 (19.60)	72.79 (21.87)	75.51 (20.20)	70.15 (20.35)	69.65 (24.37)	71.55 (21.52)
English	97.20 (4.42)	95.97 (7.54)	97.08 (6.03)	95.55 (9.86)	96.79 (6.67)	97.25 (10.08)
Punjabi	61.25 (24.50)	61.66 (23.01)	63.85 (23.45)	62.90 (25.78)	65.86 (25.68)	60.40 (23.19)
Tagalog	84.35 (13.68)	86.84 (18.56)	84.01 (16.87)	83.05 (15.38)	86.29 (16.71)	84.99 (18.19)
Accentedness						
Arabic	71.25 (12.24)	76.56 (14.85)	72.20 (20.68)	79.90 (12.51)	73.12 (20.16)	72.43 (17.05)
Mandarin	55.30 (19.63)	56.99 (24.61)	51.42 (23.55)	65.20 (21.98)	62.73 (24.79)	59.45 (21.78)
English	7.00 (8.10)	6.22 (9.65)	5.10 (8.24)	2.95 (4.71)	5.56 (8.22)	5.65 (9.53)
Punjabi	80.40 (9.21)	78.65 (14.87)	80.35 (14.22)	88.70 (8.42)	79.75 (16.73)	81.14 (18.06)
Tagalog	38.30 (26.24)	35.98 (29.24)	40.25 (27.97)	40.70 (32.76)	38.21 (32.47)	38.85 (27.64)

For comprehensibility, as summarized in Table 2, there was no significant effect of condition for any speaker group, such that the assessments in the voice–image match and mismatch conditions were comparable to those in the audio-only condition (for summary of full mixed-effects models, see Appendix D). In terms of location, Calgary and Montreal raters provided similar assessments for all speaker groups, except the Arabic speakers whom Montreal

raters evaluated as more comprehensible than Calgary raters ($p = .024$), regardless of the condition.

Table 2

Summary of Mixed-Effects Models for Comprehensibility Ratings in the Voice–Image Match and Mismatch Conditions by Location and Speaker Group

Speakers	Voice–image match			Voice–image mismatch			Location		
	<i>Estimate</i>	<i>t</i>	<i>p</i>	<i>Estimate</i>	<i>t</i>	<i>p</i>	<i>Estimate</i>	<i>t</i>	<i>p</i>
Arabic	2.51	0.57	.571	2.05	0.47	.643	5.35	2.27	.024
Mandarin	–0.23	–0.05	.959	2.08	0.46	.646	–3.48	–1.50	.135
English	0.01	0.01	.997	0.79	0.46	.644	0.35	0.42	.677
Tagalog	2.86	0.91	.364	0.80	0.26	.799	0.09	0.05	.957
Punjabi	1.69	0.34	.735	0.05	0.01	.992	0.47	0.18	.854

Note. The baseline (audio-only) condition and the Calgary location were set as reference levels. Full statistical output appears in Appendix D.

For accentedness, as summarized in Table 3, there was also no significant effect of condition for any speaker group, such that the assessments in the voice–image match and mismatch conditions were comparable to those in the audio-only condition (for full mixed-effects models, see Appendix D). In terms of location, Calgary and Montreal raters provided similar assessments for all speaker groups, except the Mandarin speakers whom Montreal raters evaluated as more accented than Calgary raters ($p = .003$), regardless of the condition.

Table 3

Summary of Mixed-Effects Models for Accentedness Ratings in the Voice–Image Match and Mismatch Conditions by Location and Speaker Group

Speakers	Voice–image match			Voice–image mismatch			Location		
	<i>Estimate</i>	<i>t</i>	<i>p</i>	<i>Estimate</i>	<i>t</i>	<i>p</i>	<i>Estimate</i>	<i>t</i>	<i>p</i>
Arabic	–0.73	–0.19	.846	–3.31	–0.88	.381	–0.95	–0.52	.606
Mandarin	–0.39	–0.09	.931	–4.81	–1.06	.293	7.12	3.00	.003
English	0.92	0.50	.617	0.40	0.22	.828	–0.36	–0.38	.702
Tagalog	–2.41	–0.65	.527	0.05	0.01	.989	0.60	0.29	.771
Punjabi	–5.35	–1.57	.118	–3.81	–1.12	.265	1.46	0.87	.386

Note. The baseline (audio-only) condition and the Calgary location were set as reference levels. Full statistical output appears in Appendix D.

Because raters evaluated all speaker groups similarly regardless of the condition in which the speech was presented, in subsequent analyses, condition was not considered further. Instead, subsequent modeling explored if raters differed in their evaluations of the five speaker groups as a function of context, since minor differences across location emerged in initial analyses (see Tables 2 and 3). In these analyses, speaker group (Arabic, Mandarin, Punjabi, Tagalog, North American English) and location (Calgary, Montreal) served as fixed effects, and an intercept for rater and an intercept for speaker were again entered as random effects (see Appendix D for full model output). Location did not emerge as a significant predictor of comprehensibility ratings (*Estimate* = 0.59, 95% CI = [–2.17, 3.35], *p* = .676), nor did location matter for accentedness ratings (*Estimate* = 1.60, 95% CI = [–0.73, 3.94], *p* = .179). However, speaker group emerged as

a significant predictor for both comprehensibility ($Estimate > 7.37, p < .001$) and accentedness ($Estimate > 6.70, p < .001$).

To understand relative differences in rater evaluations of the speaker groups, between-group contrasts were explored through analyses of estimated marginal mean differences (EMM_{diff}) using the Satterthwaite method with Tukey adjustment for multiple comparisons (see Appendix D for full summary). In terms of comprehensibility (illustrated in Figure 1), the Arabic speakers were rated similarly to Punjabi speakers ($EMM_{diff} = 2.05, 95\% CI = [-1.71, 5.81]$), receiving the lowest comprehensibility ratings. The Mandarin speakers were rated as more comprehensible than both these groups ($EMM_{diff} > 7.37, 95\% CI = [3.61, 13.19]$), and the Tagalog speakers were rated higher than the Mandarin speakers ($EMM_{diff} = 13.06, 95\% CI = [9.29, 16.82]$). Finally, the English speakers received the highest ratings of all ($EMM_{diff} = 11.40, 95\% CI = [7.64, 15.16]$). Thus, for comprehensibility, the ranking was Arabic = Punjabi < Mandarin < Tagalog < English.

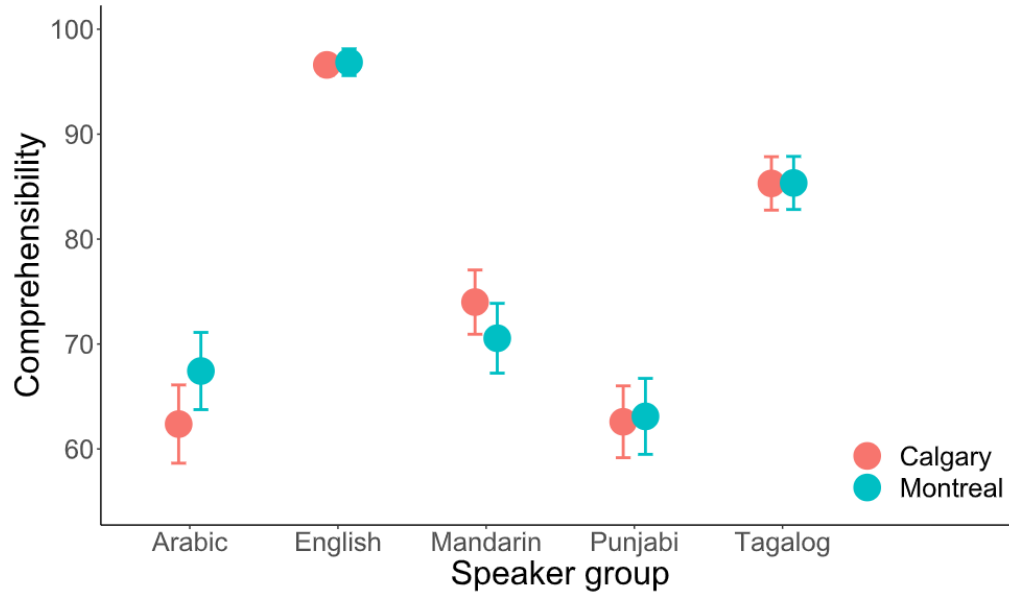


Figure 1

Comprehensibility Ratings by Speaker Group and Location (Calgary vs. Montreal), with Brackets Enclosing 95% Confidence Intervals

In terms of accentedness (illustrated in Figure 2), the Punjabi speakers were rated the most accented, followed by the Arabic speakers ($EMM_{diff} = 6.70$, 95% CI = [2.58, 10.82]), the Mandarin speakers ($EMM_{diff} = 15.84$, 95% CI = [11.73, 19.96]), the Tagalog speakers ($EMM_{diff} = 19.48$, 95% CI = [15.37, 23.60]), and finally the English speakers, who received the lowest (least accented) ratings ($EMM_{diff} = 32.89$, 95% CI = [28.78, 37.01]). Therefore, for accentedness, the ranking was: English < Tagalog < Mandarin < Arabic < Punjabi.

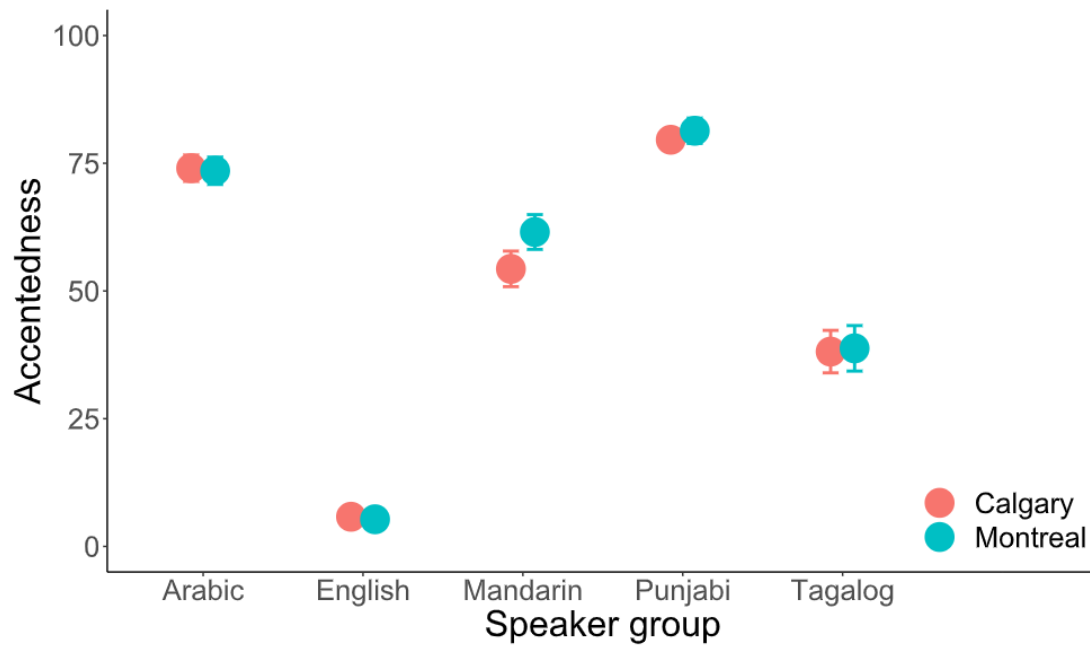


Figure 2

Accentedness Ratings by Speaker Group and Location (Calgary vs. Montreal), with Brackets Enclosing 95% Confidence Intervals

Attitudes About Immigration

The second research question asked if there were any differences between Montreal and Calgary raters' views about immigration. Responses to the three composite measures from the social attitudes questionnaire (summarized in Table 4) suggested that all raters generally agreed that individual characteristics and contributions of immigrants were moderately important, that their race and religion were relatively unimportant, and that they would generally not mind if an immigrant occupied prominent roles in society.

Table 4

Means (Standard Deviations) for Raters' Responses to the Social Attitudes Questionnaire by Location (0–100 Scale)

Variable	Calgary	Montreal
Individual characteristics	53.52 (20.75)	60.28 (23.87)
Race and religion	3.79 (10.79)	4.49 (12.86)
Roles	94.05 (11.14)	94.35 (11.04)

To determine if there were between-context differences in raters' opinions about immigration, their questionnaire responses were analyzed through mixed-effects models where location (Calgary, Montreal) served as a fixed effect, and an intercept for rater was included as a random effect. As shown in Table 5, there were significant differences between Calgary and Montreal raters in their responses to the three measures, although the 95% CI for the mean difference nearly crossed zero for the question regarding immigrant roles. Compared to Calgary raters, Montreal raters placed greater importance on desired immigrant characteristics and immigrants' race and religion and expressed marginally stronger acceptance of immigrants playing prominent roles in everyday life.

Table 5*Summary of Mixed-Effects Models for Raters' Responses to the Social Attitudes Questionnaire*

Variable	Calgary–Montreal comparison				
	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
Individual characteristics	6.77	0.23	[6.22, 7.33]	29.28	< .001
Race and religion	0.69	0.12	[0.40, 0.98]	5.63	< .001
Roles	0.31	0.11	[0.03, 0.58]	2.67	.01

Note. The Calgary location was set as a reference level.

Relationship Between Speech Ratings and Attitudes About Immigration

The final research question asked if there is a relationship between raters' attitudes toward immigrants and their speech ratings. The results of Pearson correlations, which explored associations between raters' responses to the three composite measures from the social attitudes questionnaire and their ratings of comprehensibility and accentedness (separately per location, regardless of the visual condition), are summarized in Table 6. Only two associations reached or surpassed the .25 benchmark for a weak effect (Plonsky & Oswald, 2014), and only for the Calgary raters regarding comprehensibility. In that case, stronger acceptance of immigrants in prominent societal roles was associated with higher ratings for the North American English ($r = .26$) and Punjabi ($r = .25$) speakers.

Table 6

Pearson Correlations Between Raters' Speech Ratings and Their Responses to the Social Attitudes Questionnaire

Variable	Arabic	Mandarin	English	Punjabi	Tagalog
Calgary					
Comprehensibility					
Individual characteristics	-0.12	0.00	-0.05	-0.22	-0.21
Race and religion	0.03	0.08	-0.21	-0.11	-0.10
Roles	-0.01	0.09	0.26	0.25	0.21
Accentedness					
Individual characteristics	0.10	-0.03	0.06	0.07	0.14
Race and religion	-0.05	-0.06	0.16	-0.02	0.14
Roles	0.04	-0.02	-0.1	-0.01	-0.02
Montreal					
Comprehensibility					
Individual characteristics	-0.11	-0.06	-0.07	-0.01	0.02
Race and religion	-0.09	-0.11	-0.34	0.01	-0.11
Roles	0.15	0.12	0.12	0.14	0.14
Accentedness					
Individual characteristics	0.11	-0.04	-0.14	0.08	-0.04
Race and religion	-0.08	0.07	0.25	-0.15	-0.10
Roles	-0.07	-0.09	-0.19	0.01	0.06

Note. Associations where $r \geq .25$ (small association) are bolded.

Discussion

This study's objective was to develop a better understanding of how English-dominant residents of Calgary and Montreal judge L2 speech and to explore potential connections between those judgments and residents' overall social attitudes toward immigrants. First, there were no significant differences found between Calgary and Montreal raters in how they rated comprehensibility and accentedness of L2 speakers. There was also no difference between ratings of speech samples presented as audio-only and with a matching or mismatching image. However, subtle differences emerged between Montreal and Calgary raters in terms of the accentedness and comprehensibility of certain language groups, as well as their general attitudes toward immigrants. Ultimately, though, there was no appreciable relationship between raters' attitudes, as revealed through questionnaire responses, and their ratings of L2 speech.

Ratings of Comprehensibility and Accentedness

Of key interest for our first research question was whether raters in two contexts differ in their judgments of L2 speaker comprehensibility and accentedness and whether ratings in either context are influenced by images (matched or mismatched) presented alongside the speech samples. No image-driven differences surfaced; speakers were not upgraded or downgraded when their speech was accompanied by an image that matched or mismatched the ethnicity implied by the speaker's accent, which runs counter to prior studies involving raters outside of Canada (e.g., Hansen et al., 2017a; Rubin & Smith, 1990). In other words, Canadian raters were not swayed by a speaker's suggested appearance, even though some images depicted members of ethnic groups (e.g., East Asian, Middle Eastern) that have recently been subject to discrimination in North America (e.g., Monaghan & Santos, 2020; Wu et al., 2020).

First, it is possible that raters may not have been able to definitively identify the ethnolinguistic group to which each image belonged, with the consequence that the visual may not have been fully successful at activating rater biases. Although care was taken to select images pre-rated as most prototypical for each group, prototypicality ratings were wide-ranging (65–84%), implying that some images may have been more representative of a given group than others. Second, at least some of the discrepancy in findings might also be explained by image variability, in the sense that congruent and incongruent voice–image pairings may have reminded raters of how diverse the world can be, and that speech and appearance do not necessarily go hand in hand. In the case of Rubin (1992), Hu and Lindemann (2009), and more recently, Ghanem and Kang (2021), the photo guises were of either White or East Asian faces. For guise experiments involving only two ethnicities (i.e., one in-group, one out-group), it would be easy for a rater to be influenced, as pairings of an East Asian face with native English speech, for instance, would not test the boundaries of pairings that are typically encountered in daily life. However, included in this study were images of White, East Asian, Southeast Asian, Indian, and Middle Eastern individuals. With five matches (e.g., a Southeast Asian face with a Tagalog speaker) and five mismatches (e.g., a Middle Eastern face with an East Asian speaker) for a total of 10 pairings, half of which also suggested religious affiliation, it is possible that expectations were violated to the point that raters chose to disregard the images as noise, at least in some cases. In fact, image and speech mismatches have been shown to lead to more effortful processing (Hansen et al., 2017b), and in the face of an incongruent pairing, an increased processing effort might have led to sharpened attention toward the speech sample itself (Strauss & Francis, 2017).

Finally, the broader sociolinguistic context of Canada, with its national emphasis on multilingualism (Milton, 2006) and multiculturalism (Brosseau & Dewing, 2018), might also play a role in mitigating the effect of visual ethnic information. For example, Kutlu et al. (2022) studied two bilingual environments, Montreal (English–French) and Gainesville, Florida (English–Spanish), presenting listeners from these locations with American, British, and Indian English speech samples alongside either White or South Asian female faces and asking them to transcribe the speech heard. Montrealers not only provided more accurate transcription across conditions (with no effect of image on intelligibility) but their ratings of speaker accentedness were also unaffected by visual information, unlike for the Gainesville listeners. These results were attributed to differences in how the two multilingual environments function, where multilingualism is seen as a “resource” in Canada and a “deficit” (or sign of otherness) in the US (p. 4). Just as listeners in Kutlu et al.’s study, Calgary and Montreal raters may have been largely unaffected by speaker ethnicity, as implied through images, because ideas of multilingualism and multiculturalism to various degrees have been at the core of the Canadian identity since the 1970s (Charland, 1977).

There were few Calgary–Montreal differences in ratings assigned to specific language groups, but the two differences that did emerge likely reflected rater exposure and familiarity. For instance, Montrealers gave higher comprehensibility ratings to Arabic speakers in all conditions, which aligns with Arabic’s status as the top immigrant language spoken in Montreal, a city that hosts more Arabic speakers than any other in Canada. Montrealers also assigned harsher accentedness ratings to Mandarin speakers in all conditions than did raters in Calgary, which also aligns with exposure, as Mandarin speakers account for 7.6% of the population in Calgary and only for 3.5% in Montreal (Statistics Canada, 2017).

Beyond these (minor) differences, raters in both contexts agreed about the comprehensibility and accentedness rankings of the speakers. They found the native English speakers the most comprehensible and the least accented, followed by the Tagalog and Mandarin speakers. Perceived least comprehensible and most accented overall were the Arabic and Punjabi speakers, with similar ratings for comprehensibility and slighter harsher assessments of accentedness for Punjabi than Arabic speakers. These findings align well with previous work on accent categorization. For example, Dragojevic and Goatley-Soan (2022) recently explored American listeners' attitudes toward L2 English speakers from nine different ethnolinguistic groups. Although listeners were generally poor at identifying each ethnolinguistic group based on their accent (mean accuracy of 36.8%), a clear pattern of preferred and non-preferred categorization emerged, such that some L2 groups were more likely to be categorized as "foreign" than others. For instance, European-accented speakers (e.g., French, German) were perceived more favourably in status (e.g., intelligence, competence) and solidarity (e.g., warmth, friendliness) than speakers of more stigmatized languages (e.g., Farsi, Vietnamese), with Arabic and Hindi speakers particularly downgraded in these evaluations. Moreover, comprehensibility ratings patterned with social attitudes, where higher comprehensibility was associated with more favourable status and solidarity ratings—with coefficients as high as .93—implying that difficulty understanding a foreign-accented speaker can not only trigger low social evaluations, but low social ratings can also aggravate comprehensibility issues. Thus, although Montreal and Calgary raters aligned in their assessments of L2 speakers from five ethnolinguistic groups, their ratings mirrored the majority group's view of these speakers as out-group members whose speech patterns dictate their status on the hierarchy of preferred versus dispreferred foreign accents.

Social Attitudes About Immigrants

The second research question asked whether residents of Montreal and Calgary differ in their social attitudes toward immigrants. Raters in the two contexts generally agreed that immigrants' individual characteristics and contributions are somewhat important and that their religion and race are not very important. There was also high acceptance of immigrants in roles that could hypothetically affect raters' lives in major ways, such as being responsible for their medical care or becoming a part of their family. This agreement, coupled with no major differences in speech ratings, might further reflect the two cities' shared reputation as welcoming communities for immigrants (Guo & Guo, 2016; Scott, 2020). Nevertheless, there were a few differences between the two contexts, with Montrealers placing higher importance on immigrants' individual characteristics, race, and religion than Calgarians, but being more accepting of immigrants in key societal roles. Given that English-dominant Montrealers have experienced firsthand the importance of French in all domains of public life, including employment (Bourhis, 2019), their expectations of immigrants in these areas might be higher than those of Calgarians, who are able to use English in all of their daily interactions and who are not in direct competition with a linguistic majority for vital resources. Montrealers also placed greater importance on immigrants being White and Christian than Calgarians, although raters in both contexts scored relatively low on these factors. This difference is best explained, perhaps, by demographics (Statistics Canada, 2017), as Montreal hosts a higher Christian population than Calgary (65.8 vs. 54.9%), and fewer Montrealers identify as visible minorities than Calgarians (20.3 vs. 33.3%).

The final (marginal) difference regarding Montrealers being more accepting than Calgarians of immigrants in prominent societal roles likely reflects Montreal's reputation as a

highly integrated ethnolinguistic environment (Silverstein, 2015). In other words, daily multicultural interaction across personal and professional contexts is familiar to most Montreal residents. In contrast, while Calgary hosts a more ethnolinguistically diverse community overall, immigrants vary in how they interact with native residents, as illustrated, for instance, by the firmly defined ethnic enclaves of South and East Asian immigrants (Agrawal & Kurtz, 2018) and the German community, which has experienced more seamless integration due in part to cultural, religious, and linguistic assimilation (Grenke, 2018). These minor differences notwithstanding, raters in both cities appeared to generally welcome the integration of immigrants into their personal and professional lives.

Relationship Between Social Attitudes and Speech Ratings

The final objective of this study was to explore possible relationships between raters' assessments of L2 speakers and their responses to explicit questions about immigrants on the social attitudes questionnaire. No relationships emerged, with the exception of one weak effect regarding the acceptance of immigrants in prominent societal roles and higher comprehensibility ratings for the North American English and Punjabi speakers, such that stronger explicitly held beliefs were associated with more comprehensible speech by these speaker groups. Rather than speculate as to why this weak effect emerged, it would be more prudent to discuss the general absence of associations between social attitudes and speech ratings, particularly considering that this finding runs counter to prior research (Kang & Rubin, 2009; Simon et al., 2022; Taylor Reid et al., 2019).

One explanation for the lack of relationship between speech ratings and social attitudes might be that in-the-moment reactions to the way a person looks or sounds is largely independent from attitudes expressed when asked directly about biases. Put differently, what raters do and

what they say are two different matters. Explicit (conscious) attitudes, such as those expressed on a questionnaire, are a result of fully reasoned, deliberate opinions. These attitudes are openly displayed, but they can change over time through exposure to new experiences or information, which places these attitudes largely within an individual's control (Wilson et al., 2000). However, implicit or at least less consciously held attitudes—including those that might be elicited through a matched-guise rating procedure—are likely automatic, unintentional judgments formulated with little basis or deliberation. They are typically activated by the presence of an unknown precipitate and, without conscious connection between these attitudes and intentional behaviour, they cannot be controlled (Dovidio et al., 2002; Narayan, 2019). Although it would be difficult to argue that speech ratings (with or without image pairings) are completely implicit, they primarily draw on raters' intuitive reactions rather than their carefully reasoned responses.

In a real-world example of how implicit and explicit biases might differ, Green et al. (2007) asked cardiovascular medical residents in Atlanta and Boston to participate in implicit association testing, which drew on immediate reactions to images alongside a brief patient case scenario, and to respond to a social attitudes questionnaire. While there was no evidence of explicit bias, not only did respondents show an implicit preference for treating White over Black patients, they also exhibited stereotyping of Black patients as less cooperative. In another study involving implicit association testing through audio cues and measurement of attitudes via a self-report questionnaire, Pantos and Perkins (2012) found evidence of bias in favour of US-accented speakers in participants' reactions to audio cues, but evidence of bias in favour of foreign accents in their responses to the questionnaire, highlighting the role of distinct mental processes in the

dissimilar outcome. Thus, not only do biases surface in different ways, they also do not always align.

Aside from differences in how biases manifest across different elicitation instruments, one methodological explanation for the lack of relationship between speech ratings and attitudes expressed through questionnaire responses in the current study might be the order of task presentation. In the case of Taylor Reid et al. (2019), for instance, the experiment was designed to activate existing biases *prior* to speech rating through negative comments about L2 speakers made by the researcher. In the current study, the social attitudes questionnaire was completed *after* the speech rating task, which (as argued previously) included considerable diversity of visual and auditory cues to speaker ethnicity and religion. Thus, the speech rating task may have primed raters to carefully consider or re-consider their views about immigration as a reflection of the diversity they had experienced within the rating interface. In fact, as shown through optional written comments from raters at the end of the experimental session, the social attitudes questionnaire prompted reflection, where many raters proceeded to comment on how the experience led them to evaluate their own beliefs. So, while raters would have undoubtedly approached the speech-rating task through the lens of their own experience (Baese-Berk et al., 2020), their existing biases might not have been in the foreground prior to rating.

Calgary–Montreal Comparison: What Have We Learned?

Broadly speaking, despite predicting potential differences across the two contexts, we found little substantive discrepancy in Calgary versus Montreal raters' responses. While we initially expected some differences, particularly because many Montrealers have built a strong multicultural identity through their status as bilinguals, the fact that Calgarians were similar in their judgments, despite operating more regularly in a monolingual environment, is a welcome

finding. This suggests that speaking multiple languages and navigating a multilingual environment on a daily basis, as is the case in Montreal's environment of "superdiversity" (Silverstein, 2015), might not be the only factors that influence the way immigrants are perceived in society, despite the fact that the range of such attitudes is often in alignment with levels of exposure (Bobowik et al., 2021; Lambert et al., 1966). It is further possible that the contrast between linguistic environments in Montreal and Calgary is not as stark as, for example, the Montreal versus Gainesville comparison in Kutlu et al. (2022). For instance, many Calgarians reported knowledge and use of additional languages. Although the average number of languages spoken in Calgary ($M = 1.58$) was not as great as in Montreal ($M = 2.54$), it is clear that multicultural interactions occur in both environments among these English-dominant raters.

Perhaps the larger initiatives at work to welcome immigrants within these progressive cities, both of which belong to the arguably more conservative provinces of Alberta and Quebec, put residents of Calgary and Montreal in a similar frame of mind. For example, Calgary's municipal government works closely with neighborhoods to welcome newcomers into inclusive multicultural spaces, filling in the gap between provincial policy and local action (Guo & Guo, 2016). As further illustration of Calgary's progressive mindset, the City Council recently clashed with Montreal's outer province of Quebec by voting unanimously to formally oppose Quebec's controversial Bill 21—which bans religious symbols from being worn by certain public employees—thus asserting opposition to an overt act of discrimination that has especially heavy impact on the Muslim community (Macvicar, 2019). Montreal has also demonstrated progressive thinking by regularly launching initiatives to heighten positive community support for immigrants, including a month-long public awareness campaign to encourage diversity and inclusivity in the workforce through targeted outreach to immigrant workers (Scott, 2020). Given

these examples, it would not be surprising for residents of both cities to hold similar attitudes toward immigrants.

Implications and Future Research

On the practical side, the fact that Montreal and Calgary were largely unified in their judgments of the five different varieties of English targeted in this study and were similar in their social attitudes suggests that the two cities, both progressive hubs within their more conservative provinces, might find powerful kinship in one another. A sister-city partnership, for instance, especially between two large cities on opposite sides of Canada who are united in their views of immigration, could form a powerful alliance in the creation and implementation of Canadian policies that both strengthen Canada and result in equitable treatment of immigrants. Though official sister-city alliances, or social agreements to promote cultural and economic connection, already exist for both Calgary (e.g., Quebec City and Daqing, China) and Montreal (e.g., Hiroshima, Japan and Dublin, Ireland), even an informal relationship based on similar social policy agendas between the two cities could have heavy impact in Canada. At the very least, the two cities might capitalize on their similarities as they host rapidly growing immigrant populations.

Findings also indicate that in this case, speech was a dominant enough factor to override the influence of accompanying visual information, which supports prior work on social attitudes of children and adults (Kinzler et al., 2009; Rakić et al., 2011). When pronunciation factors are the focus of these judgments, such sharp attention to speech is not problematic, but it could pose a problem if that focus on speech carries over into other assessment factors, such as judgments of individuals' personal and professional characteristics, and into decision-making in social and professional domains. For instance, credibility, a factor for which speakers have been

downgraded based simply on features of their nonnative speech (Lev-Ari & Keysar, 2010), has been shown to be a primary factor in refugee status determination (McDonald, 2014). While visual information is not necessarily essential to such judgments, a disproportionate focus on speech could certainly be detrimental. Therefore, efforts should be made toward continued research of how judgments of L2 speech influence perceptions of speakers as individuals and actions toward them.

Furthermore, while raters mentioned very little about the mismatched image and speech combinations in their post-participation feedback, a few made comments such as, “Some of the images in the voice test were unrealistic, like the East Asian wearing a Sikh turban” and “The voices mismatched the people, stereotypically,” and it is possible that these rare (but possible) combinations complicated our findings. That should not deter future exploration of these expectancy violations, however. While it is true that many people’s appearances match the way we would expect them to sound, what of the documented existence of a native Russian speaker with East Asian features; a fair-skinned, redheaded native Hindi speaker; and yes, even a Chinese man of the Sikh faith? As global mobility expands, further blurring the line between ethnicities, races, and faiths, expectancy violations will also increase. Future research might explore expectancy violations of image and L2 speech combinations and further our understanding of how disruptions of our deeply cemented stereotypes fuel our attitudes toward people who are not only viewed as out-group members, but who are further judged for defying stereotypical parameters.

From a more theoretical perspective, this study reinforces our understanding that untrained raters are consistent in their judgments of L2 speech, even across geographical contexts, and that online speech rating tasks can yield a robust sample of reliable data (Nagle &

Rehman, 2021). But it also suggests that more nuanced investigation of how social attitudes influence ratings of L2 speech might be warranted. In addition, this research yielded a positive incidental outcome. After assessing L2 speakers of various ethnicities and answering questions about immigrants in their community, many raters shared that they were grateful for the experience. One rater commented, “The questions had me challenge my own belief systems.” Another wrote, “This was eye-opening for me. I hope there aren’t any raters who mind people of other races than their own living in Canada, teaching their children or being their physician. The thought of that is painfully racist.” A third rater shared that it was a “very interesting survey that made me reflect on our society.” Studies that explore social justice can have impact long before data analysis. Simply engaging raters in reflection about social issues can help by both raising awareness of important issues and guiding individuals toward a better understanding of their own beliefs.

Limitations

The current study explored attitudes toward immigrants in two different metropolitan contexts that were, perhaps, more similar than first believed. Although the raters in the two contexts clearly differ in their frequency of English-only interactions, both Montrealers and Calgarians are residents of environments rich in ethnolinguistic diversity. Future research might explore judgments and attitudes about L2 speakers in more contrastive environments, perhaps involving replication in rural versus metropolitan contexts or contexts with little or no ethnolinguistic diversity versus richly diverse contexts and extending to comparisons made outside North America. Furthermore, given the differences between this study’s findings and other image and L2 speech guise studies, it would be important to conduct additional research involving ethnicities with limited representation in matched guise experiments (Lee & Bailey,

2022). In terms of the current study's target ethnicities, moving beyond features of pronunciation to explore ratings of L2 speakers' personal and professional characteristics as well as other contributing factors to in-group/out-group assignment of, and ultimately behaviours toward, L2 speakers in various contexts would strengthen our understanding of how immigrants are viewed in Canada.

In addition, there are several design-specific limitations worth mentioning. First of all, speech samples were limited to one male and one female per speaker group, so it was impossible to look at attitudes toward specific genders. Future research might include several gender-specific and non-binary examples. It was also impossible to look at individual pairings of speech and ethnic images, as all voice-image combinations (within each set of matching vs. mismatching pairings) were ultimately combined to maintain a sufficiently large dataset, potentially resulting in responses that cancelled each other out, at least in the mismatch condition. Finally, the goal of the study was to elicit attitudes toward immigrants. However, although all L2 speakers were, in fact, born outside of Canada, they were not specifically labeled as such for the speech rating procedure. It is therefore possible that the speakers might not have been perceived as immigrants, but rather as other less-permanent (e.g., international student, temporary resident) or even Canadian-born individuals.

Conclusion

The current study investigated native-born residents' speech judgments and social attitudes toward L2 speakers of English in two geographically separate Canadian contexts. The results bring good news: Raters were not swayed by images (matching or mismatching) of suggested ethnicity when rating the comprehensibility and accentedness of L2 speech, although raters showed a clear hierarchy in terms of their preference for L2 speakers who are less accented

and easier to understand, where some groups were at top of this continuum (Tagalog speakers) while others were ranked toward its end (Arabic and Punjabi speakers). Furthermore, Montreal and Calgary, despite their different linguistic environments, appear to be closely aligned in their ratings of L2 speech and in their outlook regarding the effects of immigration in their respective communities. These results enrich our understanding of the role of appearance in assessing L2 speakers and the role of linguistic environment in the formation of social attitudes.

Connecting Study 1 and Study 2

Study 1 showed that raters in two different sociolinguistic Canadian contexts, Montreal and Calgary, were similar in their judgments of L2 speaker comprehensibility and accentedness, but that some speaker groups were rated more favourably than others. Results also showed that raters were not affected by visual information (matching or mismatching ethnicity) that accompanied the speech samples. Overall, raters expressed generally positive (and similar) attitudes toward immigrants through the social attitudes questionnaire, but those attitudes were not associated with ratings of L2 speech.

Given that ratings of linguistic factors varied for the different speaker groups, Study 2's focus shifted to explore ratings of extralinguistic factors (status and solidarity), this time with the goal of capturing a broader Canadian perspective through the two contexts that were so closely aligned in Study 1. In addition, since the presentation of ethnic images had no effect on speech ratings, the decision was made to focus on matching ethnic images *and* religious images in Study 2, to see if a different type of visual information would influence raters. As part of this investigation, new social attitudes categories were explored.

CHAPTER 3

Study 2: Status and Solidarity Judgments of L2 Speakers in Canada as Presented in Ethnic and Religious Guise: A Multi-City Perspective

Leonardo da Vinci once wrote that “all our knowledge has its origin in our perceptions” (Richter, 1888). Indeed, our assessments of others are built on what we believe to be true about them (Nilsson, 2014) based on both our preconceived ideas about particular groups and our past experience with others (Eberhardt, 2020). While some perceptions of others might be accurate, they might also be unfair, or biased. These judgments might be especially harmful to marginalized individuals, such as those born outside a given sociolinguistic context who speak the majority language secondary to their own mother tongue (e.g., immigrants, refugees, asylum seekers). Biased judgments about second language (L2) speakers can make it even more difficult for them to secure employment and housing (Dion, 2001; Timming, 2017) and to receive general acceptance by a host community (Montreuil & Bourhis, 2001). To complicate matters, biases against people can be triggered by everything from the way they look (Rubin, 1992) to the way they sound (Sumantry & Choma, 2021) to the culture or religion they represent (Vang et al., 2019). Therefore, the broad goal of this study is to explore how native-born residents of a host nation view immigrants, specifically L2 speakers, based on the way they sound.

Sources of Bias

One framework that can help us understand how biases develop is social identity theory, which proposes that belonging is derived from group membership (Tajfel, 1959, 1972). One’s identity might develop, for example, through association with others of the same race, culture, or religion. Those who do not share the same group characteristics are then assigned to one or more out-groups. When multiple groups reside in a shared community, those who identify as members

of an in-group may prioritize protection of their own continued access to resources over the needs of an out-group, and perception of threat may increase as an out-group grows in size (Blalock, 1967). Consequently, some residents of a host community might, for instance, resist immigration, particularly when newcomers are of a different ethnicity, practice a different religion, or speak a different language, all of which may serve as justification for assignment of out-group status (Lancee & Dronkers, 2011). This resistant mindset might best be explained by group threat theory, the idea that negative attitudes are fueled by the perceived threat of another group (Blalock, 1956; Blumer, 1958), and zero-sum theory, the false perception that one group's gains inevitably result in losses to another (Von Neumann & Morgenstern, 1944). Furthermore, social identity is often driven by essentialist beliefs that group membership both defines who we are and is fixed. This inflexible view of group membership can hinder inclusivity, leading to prejudice toward others (Bastian & Haslam, 2008).

Many commonalities can serve as the basis for group membership. For example, language, as a valuable form of capital (Clark, 2006; Pendakur & Pendakur, 2002) and a distinguishing cultural characteristic (Risager, 2006), contributes to how we are perceived by others. Furthermore, biases can stem from beliefs about the entire ethnic group with which a particular language is associated (Ng, 2007), and these biases may be further influenced by individual characteristics of speakers' language, such as how strong their accent is (Kang & Yaw, 2021) or how easy they are to understand (Simon et al., 2022). Biases can also be fueled by appearance, as it provides clues about the ethnic group to which one belongs. In daily social interactions, a person's speech and appearance are often presented simultaneously, so identifying the primary catalyst for negative judgments can be complicated. Consequently, it can be difficult to decipher whether someone is being judged based on how they sound, how they look, or both.

For this reason, biases have often been explored through guise studies that pair L2 speech with images that either match or mismatch the speaker's stereotypically perceived ethnicity. For example, when speakers of American English judged accentedness of American English and Indian English speech paired with either a White or a South Asian face, they assigned less favourable ratings to speakers of Indian English in general but evaluated both varieties (American and Indian English) as more accented when presented with a South Asian versus a White face, suggesting that speech ratings were mediated by appearance (Kutlu, 2020).

To further complicate matters, negative judgments based on speech and ethnicity can be augmented by pre-existing beliefs about certain religious groups. For instance, strong national identity, particularly among Christians, is associated with a belief that non-Christian immigrants pose a symbolic threat to national culture and traditions, which translates into negative attitudes and discriminatory behaviour (McDaniel et al., 2011). In one example of how this religion-based bias manifests, Weichselbaumer (2020) explored German employers' response to applications sent with a name and photograph of the same female applicant introduced with either a German or Turkish name, with or without a headscarf. Not only did the applications showing the Turkish name receive significantly fewer responses, but applications accompanied by images of the Turkish-named woman in a headscarf received the least interest from employers. Such religious discrimination can detrimentally affect an individual's sense of belonging (Wu & Finnsdottir, 2021), mental health (Wu & Schimmele, 2021), and overall life satisfaction (Vang et al., 2019).

Status and Solidarity Judgments

Close human relationships are often built on similarities such as cultural beliefs, religion, and race (Rushton & Bons, 2005). Therefore, biases against an individual based on their linguistic, ethnic, and religious identity can interfere with societal acceptance and inclusion.

These multi-layered, socially driven biases are most often explored through the evaluative dimensions of status and solidarity. Status evaluations are typically linked to social class and include assessments of a speaker's education level, intelligence, and competence, whereas solidarity judgments are based on perception of similarities between the assessor and the speaker and include such factors as kindness, trustworthiness, and attractiveness (Fuertes et al., 2012). Historically, evaluations of L2 speaker status and solidarity have been tied to beliefs about particular language groups (Ng, 2007) as well as reactions to an individual's speech (Kang & Yaw, 2021; Simon et al., 2022), beliefs which might then be complicated by the speaker's identity as part of an ethnic or religious minority. All of these aspects of speaker identity can frame the evaluation of a person's status and solidarity by a majority language group, with more favourable evaluations often given to speakers of the standard language variety (Dragojevic & Giles, 2016).

Judgments of status and solidarity fueled by out-group assignment can be both positive and negative. For example, Bavarian-accented speakers in a German-majority context have been perceived by listeners as both less employable and more intelligent than speakers of standard German (Rakić et al., 2011). While negative biases are arguably more problematic, the mere presence of positive bias signals out-group assignment of a less favoured group or individual. In another example of how negative character and ability assumptions go hand in hand with minority group status, Frumkin and Stone (2020) explored the effects of accent, race, and age on ratings of speaker competence as witnesses to a crime. In their experiment, raters heard randomized combinations of four similar mock testimonies given by speakers of varying race (Black or White), age (young or old), and status (stereotypically prestigious or non-prestigious regional accent), then evaluated the speakers on such witness-related factors as accuracy,

credibility, and confidence. Testimonies given by speakers with lower-prestige accents and by Black speakers were rated significantly less favourably, even though the content of the testimonies was nearly identical. The same negative effect can occur when those being evaluated are simultaneously members of an in-group (e.g., professional) and an out-group (i.e., linguistic). For instance, for bilingual social workers in Australia, bilingualism was generally advantageous for successful completion of their professional responsibilities working with a diverse clientele. However, factors such as a strong accent or less-than-perfect English led to them being perceived as less credible than their co-workers and limited their opportunities for professional advancement (Harrison, 2013).

When the sound of someone's voice alludes to their status as an ethnic minority, less positive evaluations of that person's status and solidarity traits by native speakers are common. In another example of this, Mirshahidi (2016) asked native-born Iranians living in the southern US to participate in a verbal guise study in which they rated speakers of Persian from the five most common non-Persian language backgrounds in Iran (Arabic, Gilaki, Azeri Turkic, Mazandarani, and Kurdish). Those who spoke Persian with a noticeable accent were rated less favourably in status (e.g., social class, education) than those who spoke with little or no accent and, the more identifiable the accent, the more likely the speaker was to be rated negatively regarding both status and solidarity traits (e.g., honesty, friendliness), though not all solidarity ratings of noticeably accented speakers were negative. These findings align with a large body of research documenting not only the stigmatization of minority languages and the ethnicities those languages represent, but also the effect of that derogation on judgments of status and solidarity (e.g., Dragojevic & Goatley-Soan, 2022; Kang & Yaw, 2021; McKenzie, 2015; Schluter, 2021).

One sign of cultural membership tightly intertwined with ethnicity is religion (Kim, 2011), which can inform social reactions toward an individual or group. While strong religious beliefs can be associated with discriminatory attitudes (Cantone et al., 2019; Cantone & Wiener, 2017; Hall et al., 2010), both sharing religious identity and knowing that a person practices some form of organized religion can have a positive effect on social attitudes (Chuah et al., 2016). In Lybaert et al.'s (2022) recent investigation of the effects of speech variety, ethnicity, and religion on evaluations of an instructor at a Flemish university, 314 students listened to one of two history lectures, identical in content, given by a female speaker in either Standard Dutch, associated with prestige in Flanders, or Colloquial Belgian Dutch, a less prestigious, social dialect. Listeners experienced the lecture in one of many conditions in which suggested ethnicity was manipulated through various combinations of the speaker's name (either Flemish or Maghrebi) and photo in which she was shown either with or without a hijab. The standard-speaking female shown wearing a hijab was rated significantly higher for professionalism than the speaker in the non-standard and non-religious guises, likely because a hijab is a distinct marker of foreignness and high proficiency in the standard language variety signaled motivation to integrate. In this case, a symbol of foreignness, rather than suggestion of religion itself, might have been responsible for more favourable ratings, but this finding is in line with other examples of religion's positive effect on judgments of status and solidarity, including cooperativeness (Brennan & London, 2001) and trustworthiness (Tan & Vogel, 2008).

The Current Study

Prior research has established that language, ethnicity, and religion can all play a role in status and solidarity judgments, and these judgments largely determine how society values an individual (Locke, 2003). For example, such evaluations measure how a person will contribute to

society in a practical sense (e.g., education, competence) as well as how they will interact with others from an emotional perspective (e.g., friendliness, trustworthiness). Vulnerable populations, such as L2 speakers, can be particularly affected by unfair assessments of their character and abilities. Therefore, it is important to fully explore what contributes to such biases across various sociolinguistic contexts. Furthermore, given the collective importance of religious, ethnic, and linguistic identity in shaping an individual, a sharpened understanding of how those identities contribute to a person's perceived value is crucial, particularly for minority group members, who have the most to lose from being misjudged. While listener-based evaluations have proven to be an effective way to investigate these issues, and status and solidarity research is plentiful, studies that explore the role of appearance in judgments of L2 speakers are limited, and religion is rarely introduced as a variable, particularly alongside investigations of related ethnic biases. Furthermore, the simultaneous presentation of auditory and visual information in studies that investigate status and solidarity is scarce (Lu & Gnevsheva, 2021).

Given the potential limitations that can be placed on an immigrant's ability to integrate into society and to fairly access such vital resources as employment and housing, it is imperative that we have a solid understanding of how judgments against a person's character and abilities are formed, and to which factors of an immigrant's identity they are linked. Therefore, the goal of this study was to explore how native-born residents of Canada evaluate speakers of various immigrant L2s in terms of status and solidarity, based on their speech as well as suggestions of their ethnicity and religious affiliation. Additionally, this research aimed to capture general social attitudes toward immigrants in terms of their roles in society, their civic and political impact, and

to explore relationships between those attitudes and speech-based ratings of specific L2 speaker groups.

One way to thoroughly investigate how biases against L2 speakers affect judgments of their status and solidarity in Canada is through combined investigation of two representative environments: one East and one West. Montreal and Calgary, which represent two classic Canadian sociolinguistic contexts—predominantly bilingual and predominantly multilingual—are suitable litmus-test environments for pooling data that can provide solid insight about social attitudes in Canada overall, especially considering each province’s status among the top four destinations for immigrants to Canada (Statistics Canada, 2017). Similar in both general population (Montreal 1.8 million, Calgary 1.3 million) and immigrant population (Calgary 29%, Montreal 23%), both cities boast strong cultural diversity represented by over 120 different communities in Montreal and over 240 in Calgary (Statistics Canada, 2017). With diversity comes threat, however, and prior research throughout Canada has revealed the presence of negative attitudes toward immigrants based on language, ethnicity, or religion (Esses, 2021; Godley, 2018; Nangia, 2013; Schissel et al., 1989). Some of this negativity may arise from increasing competition for jobs or native-born residents’ perceived loss of national or regional identity (Danso & Grant, 2000; Montreuil et al., 2004), but feelings of threat can also arise from immigrants not learning the majority language (Gravelle, 2018). In some cases, these biases have taken center stage in Canada. One example of this can be found in Quebec’s controversial Bill 21—which bans religious symbols from being worn by some public employees—an overt act of discrimination that has especially heavy impact on the Muslim community (Macvicar, 2019).

To gain a multi-city perspective of potential biases, 340 native-born residents of Calgary and Montreal were presented with speech samples recorded by recent immigrants from four

different linguistic backgrounds, representing either the top or fastest growing immigrant languages in common between the two cities (Arabic, Tagalog, Mandarin, and Punjabi), with North American English used for comparison. Participants evaluated the audio recordings for intelligence, friendliness, and trustworthiness (among other dimensions not included in this study). These traits were chosen for their relevance to key aspects of immigrant acceptance: intelligence as a measure of immigrants' ability to contribute knowledge to and make informed decisions about society (Boyd & Alboim, 2012; Boyd & Schellenberg, 2007), friendliness as an index of their potential to integrate and form bonds with native-born residents (George & Selimos, 2019), and trustworthiness as a metric of their suitability for prominent roles in society, such as in healthcare and education (Wheeler et al., 2019). Sometimes the target speech samples were presented alone, and sometimes accompanied by images that matched the speaker's ethnicity, as in typical photo guise experiments (e.g., Kang & Yaw, 2021; McGowan, 2015), with or without an item of clothing (i.e., hijab, turban) that suggested religious affiliation. This design was implemented so that speech ratings could be explored in situations where the image prompted potential biases based on the speaker's ethnic and religious identity, relative to the no-image condition. The same participants also completed an extensive social attitudes questionnaire targeting their attitudes toward roles immigrants occupy in society and the importance of their civic and political impact. Of key interest was participants' speech ratings of various L2 speaker groups, their responses to the social attitudes questionnaire, and a potential relationship between these measures. The following research questions were explored:

1. How do native-born Canadian residents rate L2 speakers' intelligence, friendliness, and trustworthiness when speech samples are presented in three conditions (audio-only, audio paired with a nonreligious image, and audio paired with a religious image)?

2. How do native-born Canadian residents view L2 speakers in terms of their roles in society, civic impact, and political impact, and how do these attitudes relate to ratings of L2 speakers' intelligence, friendliness, and trustworthiness?

Based on the small body of prior research that suggests an ethnic or religious image effect on status and solidarity judgments (e.g., Lu & Gnevsheva, 2021; Lybaert et al., 2022) and on the larger body of matched guise research that has found evidence of negative bias toward L2 speakers activated by a bias-charged ethnic image (e.g., Kang & Rubin, 2009; Kutlu, 2020), it was expected that ratings would differ for the same speech sample when presented as audio-only versus with an image. If so, it would be clear that appearance was activating a response that surpassed judgments of the speech alone (Lu & Gnevsheva, 2021). Furthermore, given the two cities' Christian majority populations, with 55% in Calgary and 66% in Montreal (Statistics Canada, 2017), the least favourable ratings of intelligence, friendliness, and trustworthiness were expected to be given when speech samples were presented with an image of a person in a hijab or turban, both representing distinctly non-Christian religions. This prediction is based on a connection between Christianity and strong national identity as well as an accompanying belief that non-Christian immigrants threaten the stability of existing culture and traditions (McDaniel et al., 2011).

For the second research question, based on threats to social identity (Tajfel, 1972), the existence of group conflict (Esses et al., 1998), and the perception of resource limitations in the areas of employment, housing, and social programs that result, for instance, in fear that sharing resources with newcomers will result in a shortage for native-born residents (Soroka & Robertson, 2010; Von Neumann & Morgenstern, 1944), it was expected that some level of negative bias toward immigrants would surface in the social attitudes questionnaire. It was further expected

that any biases against immigrants expressed through questionnaire responses would be associated with harsher intelligence, friendliness, and trustworthiness ratings of L2 speakers, with the harshest ratings coming from those who hold the strongest negative attitudes.

Method

Raters

Raters were recruited through social media, emails to local organizations, word of mouth, and participant referrals, which produced a participant group of 340 native-born residents of Canada (170 Montreal, 170 Calgary). Participants included 200 females, 131 males, and nine non-binary individuals between the ages of 20 and 60 ($M = 38.61$, $SD = 11.35$). Because rater recruitment targeted those who used English as the primary language of interaction in their daily lives, the majority of participants (310) were native speakers of English, and any L2 English speakers included in the study self-reported English dominance. Raters with L2 dominance were specifically avoided to ensure that there would be no English proficiency issues interfering with their ability to judge L2 English speech. The average length of residence was 29 years, and all had lived in their home city for at least three years prior to the study, which allowed for inclusion of both lifetime residents and those who had moved away and returned. Using a 100-point scale, participants reported the largest percentage of their daily interaction as being with native English speakers ($M_{\text{Montreal}} = 86.00$, $SD = 17.99$; $M_{\text{Calgary}} = 82.93$, $SD = 16.86$), and many reported speaking more than one language ($M_{\text{Montreal}} = 2.54$, $SD = 0.91$; $M_{\text{Calgary}} = 1.58$, $SD = 0.90$). However, self-reported L2 proficiency varied among these participants ($M_{\text{Montreal}} = 84.60$, $SD = 14.56$; $M_{\text{Calgary}} = 61.27$, $SD = 33.78$) on a 100-point scale (0 = “can say a few words,” 100 = “completely fluent”).

Materials

Speech Samples and Photo Guises

Speakers who represented typical speech for each of the five target language backgrounds (Arabic, Mandarin, Punjabi, Tagalog, and North American English) were recruited to participate in a mock online job interview with an experienced interviewer. These target languages were chosen to reflect the three most prominent and/or fastest growing immigrant language backgrounds in the two cities. The 10 Montrealers—one male and one female representing each group—were all born outside of Canada, with the exception of the two native speakers of North American English, whose speech was used as a baseline. The L2 speakers self-reported intermediate to advanced English speaking proficiency, and all spoke with noticeable accents consistent with their native language. Though the full interview lasted 20–30 minutes, a brief segment (30–45 seconds) in which speakers discussed how they respond to constructive criticism was chosen for the speech rating task. The decision of which response to include was primarily based on clear identification of the responses as relating to the same topic. Participants received feedback on their interviewing skills and monetary compensation.

To select the 10 target images to go along with the speech samples, a set of 38 faces—including at least 3 male and 3 female of each target ethnicity—was pre-rated by 10 Montrealers between the ages of 24 and 55 ($M = 36.80$, $SD = 11.43$). The set included images of individuals between the ages of 24 and 32 ($M = 27.08$, $SD = 2.40$) chosen from the Chicago Face Database (Ma et al., 2015), the Bogazici Face Database (Saribay et al., 2018), the London Face Database (DeBruine & Jones, 2017), and from images of several volunteers recruited by the researcher. All images were chosen based on the model's self-reported ethnicity. To ensure a unified set with minimal distractors, the images were simple: plain hairstyles, no tattoos, and minimal make-

up. In addition, a Photoshop professional adjusted each image to show the person in the same grey t-shirt against the same neutral background (see Appendix A for examples). For the pre-rating task, participants were asked to identify the ethnicity of the person in the image from among five broad groups (see Appendix B for the full list), then rate each photo for prototypicality. The final set of 10 images (1 male and 1 female for each of the five target ethnicities) was based on highest rated ethnic prototypicality (65–84%) within each target ethnic group. To incorporate suggestions of religious affiliation, Photoshop was used to add hijabs to females and turbans to males, for a final set of 20 images (10 nonreligious and 10 religious).

Speech Ratings

The audio narratives were presented to raters in three conditions (audio-only, audio with a nonreligious image, audio with a religious image) and included one male and one female of each of the target languages (Arabic, Mandarin, Punjabi, Tagalog, and North American English). For the image conditions, each sample was paired with a photo that matched the speaker's ethnicity, with half of the images suggesting religious affiliation in that females were shown wearing a hijab, and males were shown wearing a turban. These religious symbols were chosen specifically because they represented groups (Muslim women, Sikh men) that have been historically discriminated against in Canada (e.g., Aujla-Bhullar, 2020; Williams et al., 2022). None of the photos were of the actual speakers.

To ensure that all pairings would be experienced with the same frequency, 16 survey versions were created, and 320 raters across Canada (160 Calgary, 160 Montreal) were randomly divided into groups of 20 per survey. This configuration resulted in a total of 80 target items (5 language groups \times 16 voice-image pairings), where the two speakers from each ethnic group were presented eight times with a nonreligious image and eight times with a religious image

across all surveys. In addition, 80 other voice–image pairings included male and female speakers presented with photos that did not match their ethnicity (with or without a religious head garment), but these mismatching conditions served as fillers and were not used for analysis. Because all possible combinations of voice–image pairings could not be presented in each survey (so that a given rater would not be exposed to the same speaker’s voice more than once), the 320 raters (160 per location) reacted to a random subset of five voice–image pairings, contributing 1,600 datapoints to the dataset. An additional 20 baseline raters (10 per city) completed an audio-only rating task in which all 10 speech samples (1 male and 1 female per language group) were presented in randomized order, contributing an additional 200 datapoints. Across all conditions, no rater heard the same audio or saw the same image more than once.

Raters evaluated each speech sample for status (intelligence) and solidarity (friendliness, trustworthiness) using a 100-point scale (for validation of similar scales, see Saito et al., 2017). In each case, the scale was labeled on the left (corresponding to 0) and right (corresponding to 100) to indicate judgments ranging from the speaker lacking the rated quality (e.g., “not very intelligent”) to exemplifying the rated quality (e.g., “very intelligent”). These ratings were made alongside ratings of comprehensibility and accentedness (discussed in a separate study). Each sample was played only once before raters were asked to move the scale toggle to the desired position between 0 and 100. All scales began with the toggle at the midpoint (50), and numbers to indicate each datapoint appeared as the toggle moved. To ensure continued engagement, raters

could not continue the task without at least minimally moving the toggle, even if only to return to the midpoint.

Social Attitudes Questionnaire

To explore raters' attitudes toward immigrants, the European Social Survey (2015) was adapted to create a social attitudes questionnaire (Appendix C), which invited raters to share their feelings about immigrants occupying prominent roles in society (e.g., "How much would you mind or not mind if someone from another country who is a different race or ethnic group from most Canadian people were your neighbor?") as well as about their civic impact (e.g., "Would you say that it is generally bad or good for Canada's economy that people come to live here from other countries?") and political impact (e.g., "How important is it for a country to encourage people who come and live in Canada to vote in elections?"). Participants rated the importance of each aspect using a scale from 0 to 100, with 100 representing either "very important" for the questions about individual characteristics or "not mind at all" for questions about immigrant roles in society. Additional questions (discussed in a separate study) explored rater attitudes about such topics as individual immigrant characteristics (e.g., work skills) and government policies (e.g., laws against racial or ethnic discrimination in the workplace). All scales were set at the midpoint (50) and, as raters clicked and dragged the toggle, numbers corresponding to each individual datapoint appeared above the scale. To ensure engagement with the question, some movement of each scale (even to return to 50) was required before raters could proceed to the next sample.

Procedure

All tasks, including a consent form, a background questionnaire, a social network survey (the data from which are discussed in another study), the speech rating task, and the social

attitudes questionnaire were completed in LimeSurvey. The choice to conduct the survey online was made out of practicality due to COVID-19 restrictions at the time of data collection. Nevertheless, online tools have shown high internal consistency comparable to lab-collected data (Gosling et al., 2004; Nagle, 2019). The survey session lasted approximately 30 minutes. To ensure focused listening, raters were asked to use personal headsets or complete the task in a quiet location, and they were encouraged to report any noise or interruptions that occurred during the rating session. Although the interface enabled raters to take breaks between sections, participants were not allowed to change completed sections or skip tasks, and overall progress was time-tracked, revealing that all participants completed the full session within 24 hours.

Data Analysis

To prepare the data, scalar ratings of intelligence, friendliness, and trustworthiness (0–100 scale) were compiled into a matrix, totaling 1,800 observations per variable. As shown through two-way, average-measure intraclass correlations, the raters generally demonstrated average to high consistency in evaluating intelligence (.56–.88), friendliness (.44–.84), and trustworthiness (.35–.82) across the surveys, although several low values implied variability among raters in evaluating these status and solidarity traits. Social attitudes questionnaire items (0–100 scale) were first checked to ensure consistency in directionality. As a result, responses to two items were reverse-coded. Based on both topic similarity and the patterning of responses in preliminary analyses, as confirmed by reliability indexes (Cronbach's α), similar questions (2–4 per topic) were grouped by theme, and a mean value was calculated per rater across all questions within each theme, producing a more robust set of responses per topic. For this study, three composite measures identified as most relevant to raters' acceptance of immigrants as contributing members of society were selected for analysis and incorporated into the matrix, with

the remaining data reserved for future studies. The first measure included four questions about raters' willingness to accept an immigrant as their boss, physician, in-law, or child's teacher ($\alpha = .86$). The second included three items related to immigrants' civic impact in terms of contributions to overall quality of life, contributions to society, and economic impact ($\alpha = .82$), and the third included two questions about political impact; namely, how important it is for immigrants to become citizens and vote in elections ($\alpha = .69$).

The first research question was explored through mixed-effects modeling using the `lme4` package (Bates et al., 2014) in R version 4.0.2 (R Core Team, 2020). Although there is no consensus regarding the criteria for considering statistical significance in mixed-effects modeling, with some scholars using $t > |2.00|$ as the benchmark of significance (Linck & Cunnings, 2015), we examined 95% confidence intervals (CIs) to check the statistical significance of each parameter (interval does not cross zero). To account for random variance, all models included random intercepts for participants (i.e., raters) and, where applicable, for items (i.e., speakers). Random slope models were also examined; however, the inclusion of random slopes did not improve model fit for any outcome variable, so the final models excluded random slopes. To address the second research question, Pearson correlations were carried out to examine potential relationships between raters' views about immigrants and their ratings of intelligence, friendliness, and trustworthiness.

Results

Status and Solidarity Ratings

The first research question explored how native-born Canadian residents rate L2 speaker status and solidarity traits for various ethnolinguistic groups when speech samples are presented in three visual conditions. The ratings of intelligence, friendliness, and trustworthiness

(summarized in Table 1) were first analyzed separately within each speaker group. In each analysis, condition (no image, nonreligious, religious) served as a fixed effect, and an intercept for raters and an intercept for speakers were included as random effects.

Table 1

Means (Standard Deviations) for Intelligence, Friendliness, and Trustworthiness Ratings by Speaker Group and Condition (0–100 Scale)

Speakers	Audio-only	Nonreligious	Religious
Intelligence			
Arabic	76.08 (17.00)	75.84 (19.16)	76.76 (18.00)
Mandarin	70.12 (21.65)	73.75 (18.96)	72.88 (17.89)
English	74.15 (19.86)	74.31 (19.00)	78.58 (17.68)
Punjabi	73.70 (21.39)	71.54 (18.66)	74.02 (18.19)
Tagalog	76.22 (17.53)	77.81 (16.81)	77.49 (15.73)
Friendliness			
Arabic	79.08 (13.72)	78.86 (17.84)	79.70 (17.12)
Mandarin	66.45 (22.85)	72.67 (18.33)	69.63 (18.01)
English	76.08 (17.91)	74.90 (18.75)	80.58 (16.49)
Punjabi	76.50 (19.71)	74.86 (17.04)	77.81 (17.26)
Tagalog	75.25 (19.07)	80.24 (15.63)	77.72 (16.21)
Trustworthiness			
Arabic	72.70 (18.49)	74.39 (19.43)	76.36 (17.43)
Mandarin	69.92 (22.76)	72.95 (18.65)	72.61 (18.30)
English	72.92 (20.25)	74.32 (17.53)	77.53 (17.57)
Punjabi	75.47 (20.82)	72.03 (19.12)	76.57 (17.54)
Tagalog	76.53 (17.24)	78.71 (16.81)	77.22 (16.29)

For intelligence, as summarized in Table 2, there was no significant effect of condition for any speaker group, such that the assessments in the nonreligious and religious conditions were comparable to those in the audio-only condition.

Table 2

Summary of Mixed-Effects Models for Intelligence Ratings in the Nonreligious and Religious Conditions by Speaker Group

Speakers	Nonreligious				Religious			
	<i>Estimate</i>	95% CI	<i>t</i>	<i>p</i>	<i>Estimate</i>	95% CI	<i>t</i>	<i>p</i>
Arabic	-0.24	[-7.92, 7.45]	-0.06	.952	0.68	[-7.00, 8.36]	0.17	.862
Mandarin	3.63	[-4.04, 11.29]	0.93	.355	2.76	[-4.90, 10.42]	0.71	.482
English	0.16	[-8.16, 8.49]	0.04	.970	4.43	[-3.89, 12.75]	1.04	.298
Tagalog	1.58	[-5.26, 8.43]	0.45	.650	1.26	[-5.58, 8.11]	0.36	.718
Punjabi	-2.16	[-10.02, 5.71]	-0.54	.591	0.32	[-7.54, 8.18]	0.08	.937

Note. The baseline (audio-only) condition was set as a reference level. Full statistical output appears in Appendix E.

For friendliness, as summarized in Table 3, there was no significant effect of condition for any speaker group, such that the assessments in the nonreligious and religious image conditions were comparable to those in the audio-only condition.

Table 3

Summary of Mixed-Effects Models for Friendliness Ratings in the Nonreligious and Religious Conditions by Speaker Group

Speakers	Nonreligious				Religious			
	<i>Estimate</i>	95% CI	<i>t</i>	<i>p</i>	<i>Estimate</i>	95% CI	<i>t</i>	<i>p</i>
Arabic	-0.22	[-7.64, 7.21]	-0.06	.954	0.63	[-6.80, 8.05]	0.17	.869
Mandarin	6.23	[-1.35, 13.80]	1.61	.110	3.18	[-4.39, 10.75]	0.82	.412
English	-1.18	[-8.76, 6.41]	-0.30	.761	4.51	[-3.07, 12.09]	1.17	.245
Tagalog	4.99	[-1.83, 11.82]	1.44	.151	2.47	[-4.35, 9.29]	0.71	.477
Punjabi	-1.64	[-8.92, 5.65]	-0.44	.659	1.31	[-5.97, 8.60]	0.35	.724

Note. The baseline (audio-only) condition was set as a reference level. Full statistical output appears in Appendix E.

For trustworthiness, as summarized in Table 4, there was again no significant effect of condition for any speaker group, such that the assessments in the nonreligious and religious image conditions were comparable to those in the audio-only condition.

Table 4

Summary of Mixed-Effects Models for Trustworthiness Ratings in the Nonreligious and Religious Conditions by Speaker Group

Speakers	Nonreligious				Religious			
	<i>Estimate</i>	95% CI	<i>t</i>	<i>p</i>	<i>Estimate</i>	95% CI	<i>t</i>	<i>p</i>
Arabic	1.69	[-6.06, 9.45]	0.43	.669	3.66	[-4.10, 11.42]	0.93	.356
Mandarin	3.03	[-4.92, 10.97]	0.75	.457	2.68	[-5.26, 10.62]	0.66	.509
English	1.39	[-5.98, 8.77]	0.37	.711	4.61	[-2.77, 11.98]	1.23	.221
Tagalog	2.18	[-5.13, 9.49]	0.59	.559	0.69	[-6.62, 8.01]	0.19	.852
Punjabi	-3.44	[-11.40, 4.51]	-0.85	.397	1.09	[-6.86, 9.05]	0.27	.788

Note. The baseline (audio-only) condition was set as a reference level. Full statistical output appears in Appendix E.

To understand relative differences in rater evaluations of the speaker groups (regardless of the condition in which audios were presented), between-group contrasts were explored through analyses of estimated marginal mean differences (EMM_{diff}) using the Satterthwaite method with Tukey adjustment for multiple comparisons (see Appendix E for full summary). In terms of intelligence (illustrated in Figure 1), the Arabic, Tagalog, and English speakers were all rated similarly ($EMM_{diff} < 1.30$, 95% CI = [-3.80, 2.58]) and were perceived to be more intelligent than both the Mandarin and the Punjabi speakers ($EMM_{diff} > 3.23$, 95% CI = [0.73, 7.11]), for whom there was no difference in ratings ($EMM_{diff} = 0.08$, 95% CI = [-2.43, 2.58]). Thus, for intelligence, the ranking was Arabic = Tagalog = English > Mandarin = Punjabi.

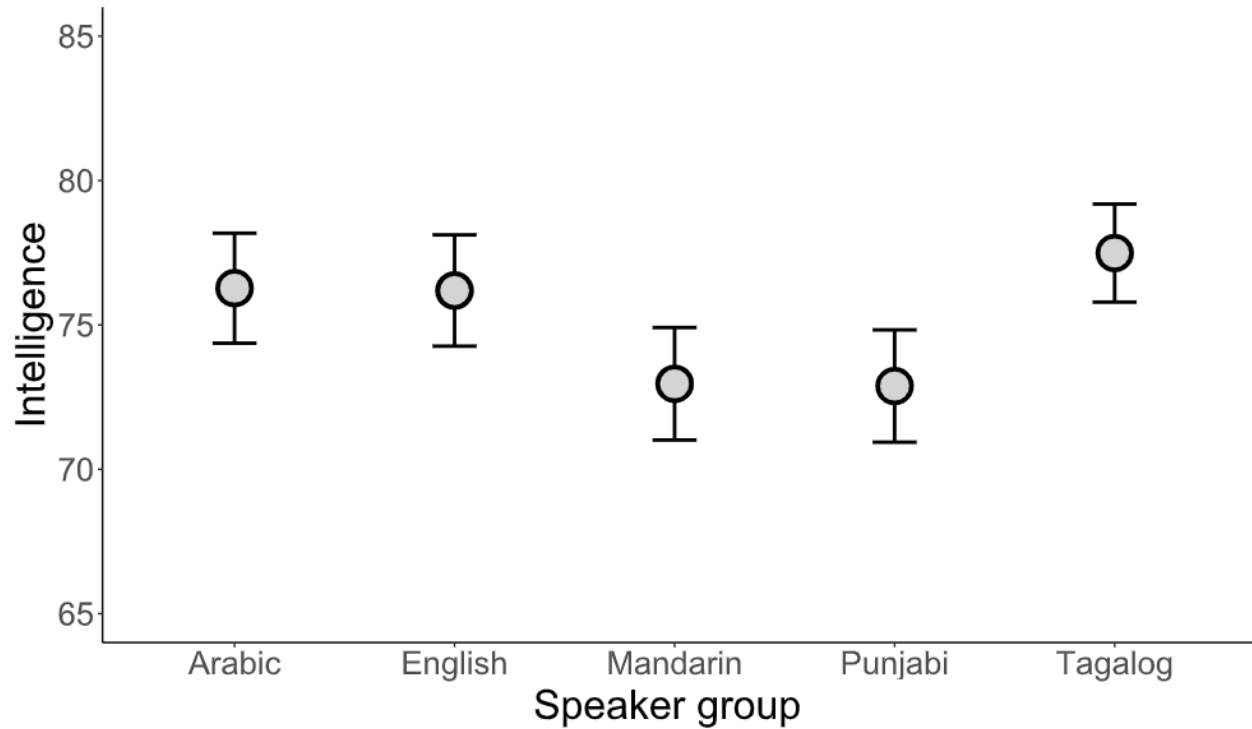


Figure 1

Intelligence Ratings by Speaker Group, with Brackets Enclosing 95% Confidence Intervals

In terms of friendliness (illustrated in Figure 2), the Arabic, English, and Tagalog speakers were rated the friendliest ($EMM_{diff} < 1.70$, 95% CI = [-3.68, 4.37]). The Punjabi speakers were rated less friendly than the Arabic speakers ($EMM_{diff} = 2.90$, 95% CI = [0.23, 5.57]), but not the Tagalog and English speakers ($EMM_{diff} < 2.21$, 95% CI = [-4.88, 3.87]). Finally, the Mandarin speakers were rated the least friendly, significantly less so than all of the other speaker groups ($EMM_{diff} > 5.72$, 95% CI = [3.06, 11.29]). Therefore, for friendliness, the ranking was: Arabic = Tagalog = English > Punjabi > Mandarin.

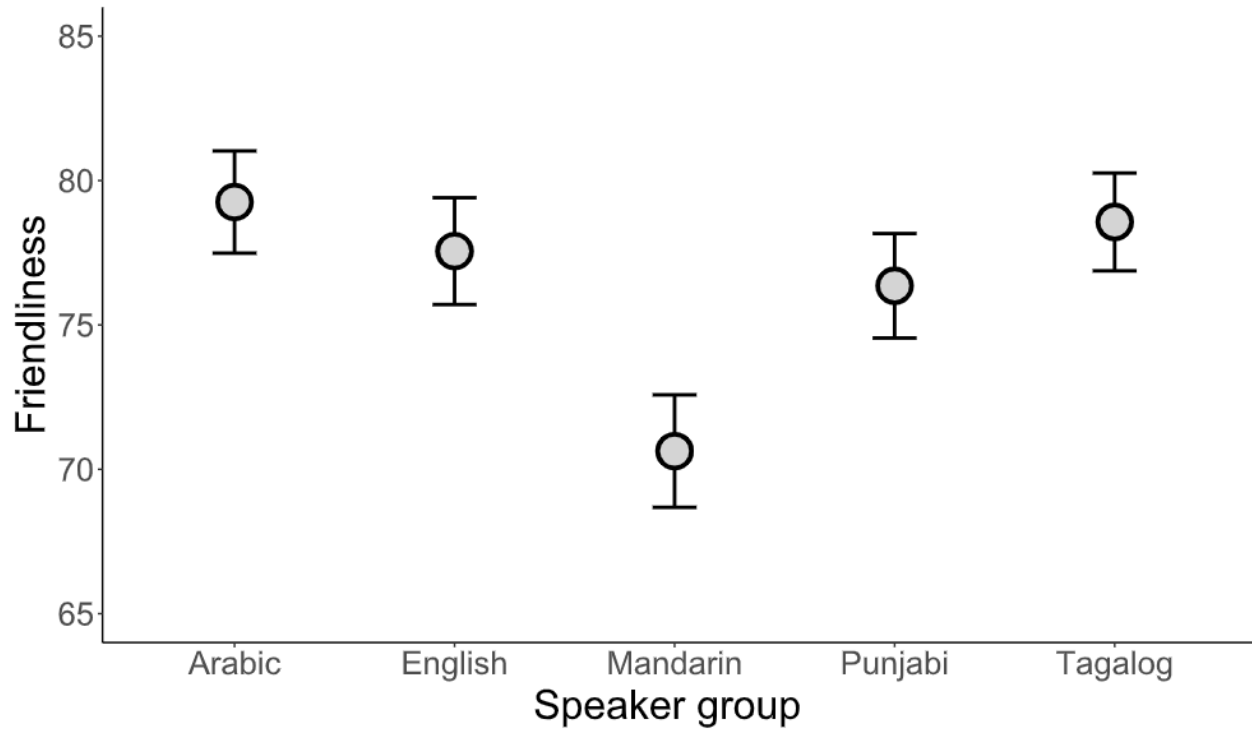


Figure 2

Friendliness Ratings by Speaker Group, with Brackets Enclosing 95% Confidence Intervals

In terms of trustworthiness (illustrated in Figure 3), the Tagalog and English speakers were rated the most trustworthy, with no difference between these groups ($EMM_{diff} = -2.21$, 95% CI = $[-4.62, 0.20]$), followed by the Arabic speakers, who were rated significantly less trustworthy than the Tagalog speakers ($EMM_{diff} = -2.72$, 95% CI = $[-5.13, -0.31]$), but not the English speakers ($EMM_{diff} = -0.51$, 95% CI = $[-2.92, 1.90]$). The Punjabi and Mandarin speakers were both rated the least trustworthy, but only the Mandarin speakers were rated significantly less trustworthy than the Tagalog, English, and Arabic speakers ($EMM_{diff} > -2.62$, 95% CI = $[-7.75, -0.21]$). Therefore, for trustworthiness, the ranking was: Tagalog = English > Arabic > Punjabi = Mandarin.

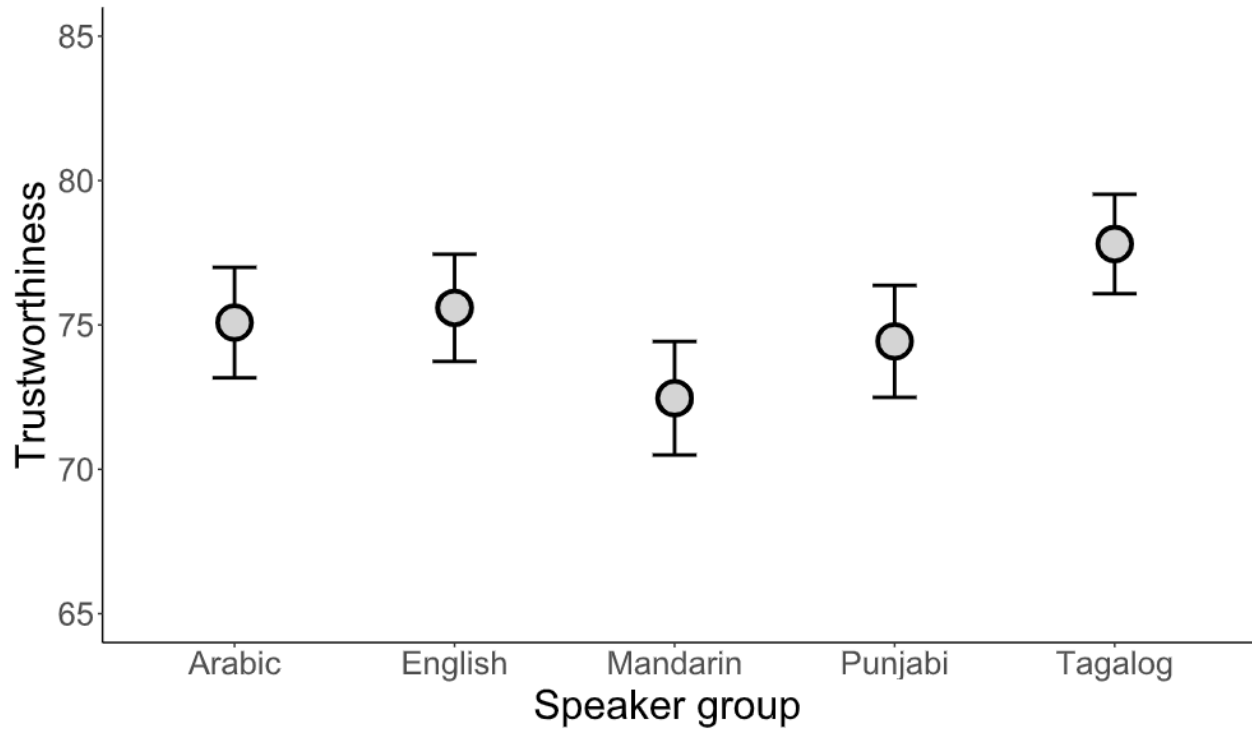


Figure 3

Trustworthiness Ratings by Speaker Group, with Brackets Enclosing 95% Confidence Intervals

Attitudes About Immigration and Relationship to Speech Ratings

The second research question asked about social attitudes held toward immigrants and the relationship between those attitudes and ratings of speaker intelligence, friendliness, and trustworthiness. First, responses to three composite measures were analyzed. Responses to the first measure showed that native-born Canadians are open to immigrants occupying prominent societal roles. On a scale of 0 (“mind a great deal”) to 100 (“not mind at all”), raters indicated high acceptance ($M = 94.17$, $SD = 13.25$) of immigrants as their neighbour, physician, in-law, or as their child’s teacher. Responses to the second measure, expressed on a scale of 0 (negative effect) to 100 (positive effect), revealed moderately high ($M = 75.74$, $SD = 22.90$) perceptions of immigrants’ civic contributions to Canada (i.e., overall quality of life, contributions to society,

and economic impact). Finally, using scalar ratings of 0 (“not very important”) to 100 (“very important”), raters placed moderately high importance ($M = 79.62$, $SD = 24.12$) on immigrants having political involvement in Canada (i.e., becoming citizens and voting in elections).

Pearson correlations (summarized in Table 5) were then used to explore potential relationships between raters’ responses to the three composite measures from the social attitudes questionnaire and their ratings of intelligence, friendliness, and trustworthiness (regardless of the visual condition). Very few associations reached or surpassed the .25 benchmark for a weak effect (Plonsky & Oswald, 2014), and those that did were primarily for intelligence and trustworthiness. Furthermore, these weak associations surfaced regarding L2 speaker groups only, not native English speakers, and concerned only perceived societal roles and civic impact, not political impact. Acceptance of immigrants in prominent societal roles was associated with higher ratings of intelligence for the Arabic ($r = .26$), Mandarin ($r = .25$), Punjabi ($r = .30$), and Tagalog ($r = .32$) speakers; friendliness for the Arabic speakers ($r = .26$); and trustworthiness for the Arabic ($r = .25$), Punjabi ($r = .28$), and Tagalog ($r = .28$) speakers. In addition, the importance of immigrants’ civic impact was associated with higher intelligence ratings for Arabic ($r = .28$), Mandarin ($r = .34$), Punjabi ($r = .32$), and Tagalog ($r = .25$) speakers as well as higher trustworthiness ratings for Arabic ($r = .25$), Mandarin ($r = .26$), and Punjabi ($r = .28$) speakers.

Table 5

Pearson Correlations Between Raters' Speech Ratings and Their Responses to the Social Attitudes Questionnaire

Variable	Arabic	Mandarin	English	Punjabi	Tagalog
Intelligence					
Roles	0.26	0.25	0.12	0.30	0.32
Civic impact	0.28	0.34	0.16	0.32	0.25
Political impact	0.00	0.14	0.14	0.14	0.10
Friendliness					
Roles	0.26	0.17	0.09	0.23	0.23
Civic impact	0.23	0.21	0.07	0.23	0.13
Political impact	0.03	0.15	0.05	0.13	0.06
Trustworthiness					
Roles	0.25	0.21	0.10	0.28	0.28
Civic impact	0.25	0.26	0.13	0.28	0.20
Political impact	0.07	0.13	0.05	0.13	0.13

Note. Associations where $r \geq .25$ (small association) are bolded.

Discussion

The goal of this study was to explore how English-dominant, native-born residents of Canada judge L2 speakers' intelligence, friendliness, and trustworthiness, and to investigate how those judgments might be related to raters' general social attitudes toward immigrants. Among the ratings of intelligence, friendliness, and trustworthiness, there were no significant differences between speech samples presented as audio-only versus with a nonreligious or religious image of a person that matched the speaker's ethnicity. However, some differences emerged in how

specific language groups are viewed in terms of their rated characteristics. A look at social attitudes through questionnaire responses revealed generally positive social attitudes toward immigrants and acknowledgment of their civic contributions, but also high expectations for their political responsibilities. Ultimately, those attitudes had weak relationships with rater judgments of L2 speakers' intelligence, friendliness, and trustworthiness.

Ratings of Intelligence, Friendliness, and Trustworthiness

Our first research question explored how native-born Canadian residents rate L2 speakers' intelligence, friendliness, and trustworthiness when speech samples are presented in three conditions (audio-only, audio paired with a nonreligious image, and audio paired with a religious image). No differences emerged based on condition. In other words, raters neither upgraded nor downgraded their ratings when the speech sample was presented alone or alongside an image, which opposes prior research showing an effect of visual information on judgments of L2 speech (e.g., Kutlu, 2020; Lu & Gnevsheva, 2021; Yi et al., 2013). Raters were not influenced by the speaker's appearance or by suggestions of religious affiliation, even though some images represented groups (e.g., Muslim women, Sikh men) that have been historically discriminated against in Canada (e.g., Aujla-Bhullar, 2020; Williams et al., 2022).

One explanation for these generally positive findings is that in terms of ethnicity, at least, the images served only as reinforcement of speaker identity, such that raters continued to base their judgments on actual speech heard. Although some photo guise experiments have shown L2 speakers to be more susceptible to employment termination than native speakers (Lippi-Green, 2012; Rubin, 1992) and to be assigned less favourable ratings of superiority and social attractiveness (Kang & Yaw, 2021), both Rakić et al. (2011) and Hansen et al. (2017a) have demonstrated that listeners might ignore visual information when presented with more salient

auditory information about a speaker's social category or in/out-group membership.

Furthermore, the effect of visual information has been primarily found for linguistic ratings of speech (e.g., comprehensibility, accentedness) rather than evaluations of speakers' competence or personality (Babel & Russell, 2015; Lee & Bailey, 2022; McGowan, 2015; Zheng & Samuel, 2017). It is therefore possible that appearance has less bearing on judgments of extralinguistic factors such as status and solidarity. It is further possible that the rating task's content (the particular response to how the speaker handles constructive criticism) and the speaker's delivery of that content (e.g., in terms of tone, volume, and pitch) were given particularly heavy consideration in ratings of status and solidarity as markers of a person's overall character, with visual information considered less relevant to the speaker's performance. Indeed, such factors have been found to play a role in a speaker's presence and individuality (Holladay & Coombs, 1994).

In terms of visual information that suggests religion, allowing the presentation of a religious symbol to influence speech ratings would imply that first, raters consider religion to be relevant to intelligence, friendliness, and trustworthiness. While some findings point to a possible connection, for instance, with respect to friendliness (Ellison, 1992) and trustworthiness (Chuah et al., 2016), support for this idea is limited. Second, any upgrading or downgrading of speaker status or solidarity based on a religious image would be an explicit demonstration of stereotyping, and Canadians might be particularly careful of this in the current political and social climate, given recent attention to such discrimination in Canada through both the media (e.g., Bartko et al., 2021; Fenn, 2021) and research (Wilkins-Laflamme, 2018; Wu & Schimmele, 2021). The weight of Canadian society's debates about religious discrimination is perhaps best illustrated by Quebec's Bill 21, a law which bans workers in the public sphere from

wearing religious symbols such as crosses and turbans and which “disproportionately impacts people who are already marginalized” (Canadian Civil Liberties Association, 2022). In addition, the raters in this study, half of whom are minority group members themselves (Anglophones in Montreal), might be reluctant to overpenalize speakers for their triple status as members of linguistic, ethnic, and religious out-groups. Finally, the absence of an image effect aligns with our understanding of Canada’s push toward inclusivity, with its national emphasis on multilingualism (Milton, 2006) and multiculturalism (Brosseau & Dewing, 2018). So, although L2 speakers are still seen as out-group members, the role of visual information in social categorization appears limited, at least in the present dataset.

Although there was no image effect on ratings, significant differences emerged in the way various speaker groups were rated for status and solidarity. Raters judged the Tagalog, Arabic, and native English speakers to be the most intelligent, followed by the Mandarin and Punjabi speakers. Similarly, raters found the Tagalog, Arabic, and native English speakers to be the friendliest, followed by the Punjabi speakers. The Mandarin speakers were rated the least friendly among all five speaker groups. A similar pattern was found for trustworthiness in that Tagalog and native English speakers were rated the most trustworthy, followed by Arabic speakers, with Punjabi and Mandarin speakers rated the least trustworthy. Notably, the patterning of results suggests that judgments of the three traits (intelligence, friendliness, trustworthiness) are likely made as part of an overall character judgment rather than as three completely separate factors. To put it another way, when speech or appearance signals a particular ethnic group, and negative attitudes about that ethnic group are held by the listener, those attitudes can translate to lower ratings across multiple status and solidarity measures (Kang & Yaw, 2021). While it is not surprising that certain L2 groups would be perceived less favourably (Dragojevic & Goatley-

Soan, 2022), what is most interesting is that Tagalog and Arabic speakers patterned with native English speakers for intelligence and friendliness, and Arabic speakers followed closely behind native English speakers and Tagalog speakers in terms of trustworthiness.

Favourable judgments across categories for native English speakers align with prior research, as their shared social identity with the raters themselves would undoubtedly result in kinship-based ratings (Tajfel, 1959, 1972). Likewise, Tagalog speakers, who represent a substantial portion of the immigrant population in both Calgary (10%) and Montreal (4%), share much in common with Anglophone Canadians. Ninety-nine percent of Canadians of Filipino origin can speak English and/or French proficiently, 41% claim English as their native tongue, and 70% speak primarily English at home (Lindsay, 2007). In addition, 97% are Christian, which echoes the religious profile of 55% of Calgarians, 66% of Montrealers, and of 63% of Canadians overall (Statistics Canada, 2017). Furthermore, although Filipinos are undoubtedly discriminated against in Canada (Mbakogu et al., 2021), they have enjoyed some advantages over other immigrant groups, including faster attainment of homeownership (Preisler, 2021). They have also been consistently acknowledged as the fastest immigrant group to contribute to Canadian society and economy after arrival (Edmonston, 2021). In sum, Tagalog speakers' language proficiency, shared religion, and positive contributions to Canada would all likely factor into highly favourable perceptions of their status and solidarity.

However, Arabic speakers also received favourable ratings across the three measures, which is particularly interesting considering that these same speakers were rated as more accented than other L2 groups by the same raters in another study, a finding that aligns with prior research comparing Arabic-accented English to other varieties (Barona, 2008). However, perceptions of accentedness and social categorization do not necessarily go hand in hand (Sa'd,

2018). Demographics of the Arab population in Canada may at least partially explain these results. First, many are native speakers of the country's two official languages, with 26% claiming English and 17% reporting French as their mother tongue. In addition, Canadians of Arab origin are twice as likely to hold a university degree as Canadians in general, are more likely to be employed, and report a strong sense of belonging through high community and political engagement (Lindsay, 2007), all of which likely contributed to raters' positive feelings toward this language group and perceptions of it as beneficial to Canada. Exposure may also have played a role (Kang & Yaw, 2021), in that half of the raters reside in Montreal, which hosts Canada's largest population of Arabic speakers (Statistics Canada, 2017).

Receiving less favourable ratings in all three categories were the Punjabi and Mandarin speakers. This outcome is surprising, as Punjabi speakers represent one of the fastest-growing immigrant groups in Canada and are the third largest immigrant language group (Statistics Canada, 2017). Furthermore, Punjabi speakers have made significant contributions to Canadian politics (Johnston, 2005), the majority are proficient in one or both official languages, and they are more likely to hold a university degree than most Canadians (Lindsay, 2007), all of which might be expected to translate to favourable ratings of status and solidarity. However, there are several social identity factors that set Punjabi speakers apart from other immigrant groups and, more importantly from native-born residents. First, Punjabi speakers have been significantly slower to acculturate in Canada than other groups with similar histories, in part due to a strong connection to their homeland (Johnston, 2005), and acculturation is often dependent on the quantity and quality of immigrants' interpersonal relationships (Jian, 2012; Sam & Berry, 1995), which are difficult to form without fully embracing the host community's culture. Furthermore, given that substantial differences in religious beliefs can interfere with acceptance (Lancee &

Dronkers, 2011; Rushton & Bons, 2005), Punjabi speakers may have been downgraded based on their primary affiliations with Sikh, Hindu, or Muslim faiths (Lindsay, 2007). While Arabic speakers might also be perceived as religious out-group members by some, the Arab-origin population in Canada is, in fact, 44% Christian.

Of the five language groups, Mandarin speakers received some of the lowest ratings for intelligence, friendliness, and trustworthiness. The intelligence ratings are particularly surprising, considering a widely held stereotype of East Asians being a “model minority” with significant contributions to society (Cheng, 1997, p. 298), especially in the science, technology, engineering, and mathematics (STEM) communities. Canadians of Chinese origin are also the largest immigrant group and are highly educated (Lindsay, 2007). However, these ratings were collected at COVID-19’s height, and research has since documented a relationship between the pandemic and a significant increase in discrimination and violence against East Asians in Canada, fueled by overall negative perceptions about East Asians (Miconi et al., 2021; Wu et al., 2020). It follows, then, that Mandarin speakers might actually *be* less friendly (understandably so) in a time of violence and threat against an entire culture.

Social Attitudes about Immigrants

The second research question explored native-born residents’ social attitudes toward immigrants in Canada. Native-born Canadians appeared to be highly accepting of immigrants in prominent societal roles (e.g., physician, neighbour), to perceive immigrants as having a generally positive effect on Canadian society and its economy, and to expect immigrants to carry political responsibility (i.e., vote and become citizens). While general acceptance of immigrants in prominent societal roles was expected, the level of acceptance (94/100) was a pleasant surprise and may be due, at least in part, to the rich diversity in both cities where participants

were recruited. In other words, Canadians in large cities have seen firsthand how immigrants contribute positively to their communities, as healthcare workers (Atanackovic & Bourgeault, 2013) and teachers (Schmidt et al., 2010), for example. Furthermore, the general perception that immigrants have a positive impact on Canada's society and economy is likely linked to Canada's attempts to prioritize immigrants who can contribute valuable skills (Chand & Tung, 2019; Lenard, 2018). Finally, a relatively high expectation of political involvement aligns with Canadian citizens' pride in both Canada's diversity and in their own citizenship (MacKinnon, 2022). This expectation is fulfilled, as the majority of immigrants who are eligible voters participate in elections (Lindsay, 2007) and eventually become citizens (Hou & Picot, 2019).

An additional objective was to explore possible associations between social attitudes and ratings of intelligence, friendliness, and trustworthiness. Although there were a handful of weak associations, no prominent relationships emerged, which runs counter to prior research (Kang & Rubin, 2009; Simon et al., 2022; Taylor Reid et al., 2019). For one, how raters respond to direct questions about their attitudes toward immigrants and what they do when judging an immigrant's speech or appearance might be two different things. Questionnaire responses are given based on reasoning and deliberation, which means that they are easy for the respondent to control (Wilson et al., 2000). Speech ratings, on the other hand, are generally intuitive, and are therefore influenced by unknown precipitates that are largely beyond a person's awareness (Dovidio et al., 2002; Narayan, 2019), which would explain a lack of strong associations between questionnaire-based attitudes and speech-driven ratings.

Of the findings that did emerge, there are three particular outcomes worth highlighting. First, associations surfaced only for the four L2 speaker groups, not the native English speakers. This is a logical outcome, as social attitudes toward immigrants would likely have little to do

with evaluations of native English speakers. Second, associations primarily concerned intelligence and trustworthiness, but not friendliness. This may be in part due to the particular social attitudes questions that were the focus of this study. While friendliness might be of some importance in a neighbor or even a teacher, it is an arguably less impactful factor for professional relationships and has little to do with a person's contributions to economy or national identity. Third, associations mostly concerned attitudes about immigrants in prominent societal roles and civic impact, but not political impact. This outcome is likely again intertwined with the nature of the evaluative dimensions. Successful professional relationships, especially high-stakes relationships between patients and their physicians, for instance, are dependent on trust (Lee & Lin, 2005); personal relationships are built on both trust and friendship (Galupo & Gonzalez, 2013); and intelligence (in all its forms) is highly valued in society and often linked to multiple measures of success, including economic outcomes (Herrnstein & Murray, 1994). In contrast, voting and citizenship have more to do with a person's national pride, sense of duty, and motivation to support a larger national identity than their individual intelligence, friendliness, or trustworthiness (Nemțoi & Ignătescu, 2014), so it is not unexpected that these personal dimensions had little association with a questionnaire-based measure of political impact.

Implications and Future Research

Although the findings of this study have showcased the welcoming identity to which many Canadians lay claim (Lund & Hira-Friesen, 2014), not all immigrant groups are viewed equally. It is clear that some language varieties enjoy higher prestige than others, likely due to pre-existing beliefs about particular ethnolinguistic groups that extend to all speakers of that language. This kind of blanket assessment of speakers can be damaging and unfair, especially when initial contact with someone of a particular language group is negative. Importantly, however, listeners are not always swayed by visual information in their assessments of L2

speech. In other words, a person's ethnicity or religion does not always influence how they are perceived by others. This is noteworthy, because judgments of a person's character or abilities have the potential to penalize a speaker for belonging to multiple minority groups. For example, a speaker might be negatively perceived based on their L2 status, their ethnicity, and their religion, piling on a string of hurdles that immigrants must overcome.

While it is difficult to know if real-world encounters between native-born Canadian residents and immigrants would yield the same results when two humans are face to face—as opposed to one interacting with a computer and the other represented only through a voice and a photograph—speech undoubtedly carries significant weight in the assessment of someone's character. Therefore, a person's L2 speaker status can affect high-stakes perceptions about them as an individual including, for instance, their employability, credibility in the courtroom, and suitability as a social acquaintance. These judgments can, in turn, affect a newcomer's ability to integrate, deprive them of vital opportunities and, most importantly, limit their chances of feeling like they belong. To complicate matters, certain language groups are judged more harshly than others. So, when a particular language group is already facing substantial discrimination based on ethnicity (e.g., Chinese people during COVID), added negativity about their language abilities can make successful integration nearly impossible. Of course, this issue is magnified during a pandemic, when all social networks are unusually strained (Rebhun, 2021).

Future research should continue to explore the influence of speech on status and solidarity judgments and, given mixed findings in prior research, the influence of ethnicity and religion. While it is clear that visual information does not always factor into rater judgments of L2 speech, it is unclear whether differences in findings across studies are due to the rater population (e.g., inclusive vs. exclusive environments), the speakers' origin (e.g., East Asian vs.

Arabic), or something else entirely. Regarding context, for instance, most prior research that has shown an effect of visual information on ratings of L2 speech has been carried out in the United States, where treatment of immigrants differs significantly from Canada in that newcomers are perceived by many to be a liability in the United States and an asset in Canada (Harell et al., 2012). Given the current study's positive findings in the Canadian context, it would be important for future research to continue to involve various combinations of Canadian raters and L2 speakers and to conduct comparative investigations between Canada and other countries.

Limitations

There are several limitations of this study worth noting. First, the current study involved participants in Montreal and Calgary, two major Canadian cities with rich diversity, multilingualism, and multiculturalism, where we would expect a greater-than-average acceptance of L2 speakers and immigrants as a whole. Therefore, it would be important to investigate native-born residents' ratings and attitudes in rural communities, in cities that host fewer immigrants, and in monolingual environments. In addition, speech samples were limited to one male and one female of each language group, so it was impossible to look at attitudes toward speakers of different genders, and language group representation was limited. Future studies should perhaps involve fewer L2 groups with more substantial representation. The content of the speech samples might also have been a limitation. First, the job interview response to handling constructive criticism is likely quite different from the more conversational speech that the raters are likely to encounter in their everyday lives. It would be worth exploring status and solidarity ratings using more informal speech samples. In addition, although the topic of each speaker's response was similar, it is impossible to disentangle individual speaker effects such as personality, confidence, and approach to the question, all of which can vary greatly among

speakers of the same language group. To put it another way, ratings might be as much about the individual speaker as the language group they represent, which is indirectly reflected in less-than-optimal reliability indexes for rater-assessed speaker traits in this study. One way to control for these factors and others (e.g., speed, fluency, volume) might be through more extensive pre-rating or through the use of several speech samples (e.g., involving personal topics or illustrating professional performances) recorded by the same speaker. Finally, the order of the tasks may have affected the overall outcome of the study. Participants were asked to rate the speech samples, then complete the social attitudes questionnaire. There are advantages to this order, because it could lead to more intuitive judgments. However, putting the social attitudes questionnaire first might prime raters, activating dormant or subconscious biases about immigrants that then influence ratings of L2 speakers.

Conclusion

The current study investigated native-born residents' status and solidarity judgments of and social attitudes toward L2 speakers of English across Canada. Raters were not swayed by the inclusion of a photo matching the speaker's ethnicity or showing religious affiliation when judging speaker intelligence, friendliness, and trustworthiness, but there were clear differences in status and solidarity judgments for various speaker groups, where some groups were rated more favourably (native English, Tagalog, and Arabic speakers), and some less so (Punjabi and Mandarin speakers). These results further our understanding of the role of speech, ethnic appearance, and religious affiliation in assessment of various L2 speaker groups and showcase generally positive attitudes toward immigrants in Canada.

Connecting Study 2 and Study 3

Study 2 showed that once again, Canadian raters were not influenced by visual information in their ratings, this time of intelligence, friendliness, and trustworthiness. However,

similar to Study 1, a clear speaker group hierarchy emerged with ratings patterning similarly across status and solidarity traits and in general alignment with the comprehensibility and accentedness ratings in Study 1. Once again, raters expressed generally positive attitudes toward immigrants through questionnaire responses, this time regarding immigrant societal roles, civic contributions, and political responsibilities, but those attitudes were not associated with speech ratings.

Recognizing a general pattern of differences across the first two studies in how raters judged speaker groups, one more investigation was undertaken using the same visual stimuli presented in Study 2, but this time for a previously unexplored rating category in a matched-guise study: judgments of citizenship status. Of key interest was whether the general pattern of speaker hierarchy that emerged for ratings of linguistic and extralinguistic factors would hold for judgments of belonging. Furthermore, since social attitudes were not found to influence ratings in Studies 1 and 2, raters in the two cities were once again divided to enable comparison of their social network exposure, with the goal of exploring the possible effects of L2 speaker interaction in two different sociolinguistic contexts on ratings of L2 speech.

CHAPTER 4

Study 3: Perceptions of L2 Speaker Belonging and the Role of Social Networks in Two Canadian Contexts

Canada's population is rapidly diversifying. In fact, by the year 2040, immigration is expected to be exclusively responsible for the nation's population growth (Vomiero, 2018). Although Canada is well known for broad acceptance of newcomers (Lund & Hira-Friesen, 2014), it is unclear whether that acceptance is based in practical tolerance—for instance, for economic purposes—or in full recognition of immigrants' belonging not only as contributors to Canada's success but also as equal citizens. Although immigrants are highly employable and contribute substantially to Canadian civic and political efforts, they still report significant discrimination based on marginalizing factors, such as their ethnic appearance, religious affiliation, and language use (Lindsay, 2007). At the root of at least some negative attitudes is the perception of threat, which can determine whether native-born residents of a given country will share their status as citizens with those against whom they compete for jobs (Raijman et al., 2008) or societal resources like medical care and housing (Ipsos, 2016), even when the native-born population is largely in favor of immigration. There may also be concern that the introduction of new ethnic and religious cultures will threaten national identity (Valentino et al., 2013). In other words, while it is one thing for native-born residents to welcome immigrants into their space for mutual benefit, it may be another to afford newcomers the same rights and privileges that residents have had since birth.

Given that native-born residents are called upon to make judgments about immigrants, most of whom are second language (L2) speakers of a community's majority language, in numerous high-stakes daily contexts—employment (Carlson & McHenry, 2006; Timming,

2017), healthcare (Halim et al., 2017), and education (Winke et al., 2013), to name a few—it is important to know how biases develop and how they affect acceptance of L2 speakers into society. Toward this goal, prior research has established that harmful biases not only exist, but also that they can be triggered by multiple factors including a person’s ethnicity, religion, and language. What remains unclear is how each of these factors influences native-born residents’ overall judgments of immigrant belonging, such as whether someone is seen as a permanent member of society, a visitor, or as someone who does not belong at all. It is further important to understand whether having a diverse social network influences these perspectives. Therefore, the overall objective of this study was to explore the link between native-born residents’ judgments of L2 speech and perceptions of immigrant belonging in Canada, in relation to residents’ social networks. Casting light on these perceptions can provide vital information about how to combat negative biases that limit host community acceptance and restrict immigrant belonging.

Perceptions and Social Bias

A growing body of evidence has illustrated how perceptions held by native-born citizens about certain ethnic groups can influence societal acceptance. For example, Scheepers et al. (2002) explored European citizens’ opposition to civil rights for immigrants through an ethnic competition theory framework and found that the more a particular ethnic group was perceived to be a threat to native-born residents of a similar socioeconomic class—as competition in the job market, for instance—the greater the resistance to extend legal rights to that group. In fact, perceived threat was found to drive societal exclusion in general. Similarly, immigrant acceptance by native-born residents of a host society can be affected by religious affiliation. For example, in a study of Muslim immigrant acceptance in Spain, Croucher et al. (2014) found that highly religious native-born citizens were more likely to harbour negative attitudes toward

Muslim immigrants than other immigrant groups, because they believed that those religious differences would reduce Muslim immigrants' motivation to adapt. Furthermore, although intergroup contact has been found to reduce cultural animosity in some cases (e.g., Dhont et al., 2014; Tropp & Pettigrew, 2005), the more Spanish Catholics had contact with Muslim immigrants in this case, the more they perceived those immigrants to be both a realistic (e.g., job-related) and symbolic (e.g., identity-related) threat.

Not surprisingly, certain immigrant groups can receive more support than others. In Canada, for instance, there is wider acceptance for young immigrants, because they are believed to place less strain on societal resources (Gravelle, 2018; Sharples & Chasteen, 2021). Similarly, certain linguistic, ethnic, and religious groups (e.g., East Asians, Christians) are often perceived to contribute more to society and to be less threatening to national identity (Lawlor & Tolley, 2017; Wong et al., 1998). Ultimately, those perceptions of how a particular immigrant group will affect society can lead to easier or more difficult paths to citizenship. Take, for example, the arguably easier path to Canadian citizenship for European immigrants than for "deviant" immigrants who do not mirror the identities of Canada's majority ethnic, religious, and linguistic groups (Winter, 2014, p. 57).

Language-Based Biases

Ethnicity and religion, often signaled by visual cues, are major factors in the marginalization of many immigrants. The nuances of language-based discrimination, however, are a bit more complicated to explore and address. For example, while a job applicant cannot be legally turned down (in most situations, at least) based on the way they look or the cultural practices in which they engage, language proficiency can be used to legally justify unsuitability in a variety of high-stakes contexts (e.g., education, employment, citizenship). In other words,

language can be used to explain access barriers that may, in fact, be driven by multiple discriminatory factors (Ng, 2007). This is decidedly tricky territory, as there is a difference between someone lacking sufficient language proficiency to effectively communicate or perform a task and someone possessing speaking characteristics that are simply perceived by the listener to be undesirable. For some listeners, the sound of speech that differs from a local variety can interfere with typical character-based assessments humans make of each other upon first introduction. Indeed, as noted by Lippi-Green (1994, p. 166), “prejudiced listeners cannot hear what a person has to say, because accent, as a mirror of social identity and a litmus test for exclusion, is more important.”

While judgments of L2 speakers that arise from social biases can be both positive and negative, those who speak the variety of the majority group often view the speech of minority groups negatively. This is because for every dominant language group, there must be one or more subordinate groups (Tajfel, 1974). Therefore, speech, a clear marker of social identity (Bourhis & Giles, 1977) carrying a wealth of identifying information (de Souza et al., 2016), makes L2 speakers an easy target for group subordination. These negative views manifest themselves among speakers of the same native language—as illustrated, for example, by English speakers’ attitudes toward speakers of African American English (Gaither et al., 2015) and Spanish speakers’ attitudes toward Puerto Rican Spanish (Fayer & Krasinski, 1987)—as well as among those who speak a L2 in a community of native speakers, such as Australian workers’ attitudes toward L2 speakers in the workplace (Harrison, 2013) and English speakers’ attitudes toward Spanish-accented speakers (Ryan et al., 1977). While even positive bias can affect the validity of speech ratings, it is negative biases that have the potential to cause the most damage to L2 speakers’ daily lives (Lippi-Green, 1994), leading to inequality, discriminatory treatment,

and outright harassment (Munro, 2003). In fact, for negative biases to arise, a mere perception of otherness is sufficient. As Lindemann (2003) has shown, language ideology might function independently of ethnic identification. She tested 39 Michigan undergraduates, who were able to distinguish native English speakers from L2 speakers of English (all Korean) in audio clips. However, the undergraduates were unable to identify the accent they heard, in that they were unclear whether the (Korean) accent they heard was indeed Korean, Chinese, Japanese, Latino, or something else. Still, the speakers were simply identified as member of an out-group: They all were rated as less intelligent and of lower status than the native speakers.

Tajfel's (1972) work presents social identity theory as a natural fit for investigation of listener attitudes toward L2 speakers. L2 speech is a clear indicator of foreign (i.e., out-group) status and often a cue to an individual's ethnicity, home country, and culture. Therefore, a L2 speaker may be viewed in a negative light simply by comparison to the listener's identifying group, and something as simple as having accented speech can uniquely disqualify a speaker from numerous opportunities (Lippi-Green, 1994). Unlike female professional musicians, for instance, who have recently been given the opportunity to audition for symphony orchestras from behind a curtain to conceal their gender, therefore allowing for unbiased judgment of their abilities (Goldin & Rouse, 2000), L2 speakers have no such method by which to conceal their otherness. Though certain ethnic identifiers like skin colour and manner of dress may be concealed in phone interviews, for example, nonnative speech persists as a conspicuous signal of minority status.

The Role of Social Networks

Social network analysis, first introduced to the field of sociology by Eisenstadt (1961), is one tool that can refine our understanding of minority group biases, and it has long been

favoured by anthropologists and social scientists as a means of understanding how relationships affect individual behaviour (McPherson et al., 2001). Social networks can not only build familiarity and lessen social distance, as with intergroup contact theory (Allport, 1954; Pettigrew, 1998), but also increase exposure to diverse speech, thereby fostering in-group receptivity to L2 speakers. In fact, since its inception, mapping of an individual's social network has been used to monitor everything from mental health (Tracy & Whittaker, 1990) and wellbeing of transplant patients (Lewis et al., 2000) to the network's role in business success as social capital (Putnam, 1993), its effect on management performance (Sparrowe et al., 2001), and its influence on conflict, the balance of power, and innovation (Schiffer & Hauck, 2010). In the past decade, however, social network theory has been used as a framework for explaining (and overcoming) at least some attitudes toward unfamiliar speech.

Social network theory in the linguistic realm posits that “people who are exposed to linguistic input from many sources should be less susceptible to the influence of new incoming linguistic input compared with people who only interact with few people” (Lev-Ari, 2018, p. 32). Thus, the theory is especially relevant for individuals in remote, rural, or predominantly monolingual communities. Furthermore, socioeconomic power and status are heavily linked to cities (Boschken, 2003; Harding & Blokland, 2014) in that they are hubs of economic activity, diversity, and employment, which makes social network theory a good opposing framework to status theories (e.g., language ideology, social identity) for examining listener attitudes among those who may be of lower prestige or themselves members of an out-group. Essentially, larger social networks lead to higher variability in interpersonal contact and, subsequently, better understanding of previously unknown (unfamiliar) speakers' accented speech (Sumner, 2011) and greater acceptance of cultural minorities in general (Tropp & Pettigrew, 2005).

The fact that social network exposure can soften attitudes toward out-group members has been well documented in research. For example, Ahmed et al. (2021) conducted a survey of Singaporean citizens that included questions about their attitudes toward Chinese-born immigrants and about their daily contact with people who differ from them in group membership (e.g., race, social class). Researchers found that discriminatory attitudes against immigrants were significantly decreased among citizens who regularly engaged in conversation with people of diverse backgrounds. The role of social networks in the formation of attitudes can also be seen through investigations of intergroup contact, even if the extension of one's personal network is brief. In one example of this, Kang et al. (2015) conducted two studies involving undergraduate students who held negative attitudes about international teaching assistants (ITAs) that resulted in reluctance to communicate. However, when the students and ITAs engaged in informal conversation and cooperative problem-solving activities, the students not only found the ITAs easier to understand, but the experience also reduced their prejudiced views (as shared through interview and questionnaire responses).

In consideration of both intergroup contact and social network perspectives, it makes sense that exposure to L2 speech might influence negative attitudes about those speakers—extending to global stereotypes about their ethnicity and religious affiliation—based on unfamiliarity and social distance, for instance. Furthermore, unlike language ideology and social identity, which are rooted in deep-seated group- and listener-centric values, an individual's social network is open to change. While the growth of social networks among those who are already immersed in high-variability multilingual environments is arguably limited, the effect of varying amounts of daily L2 interaction on social attitudes is nevertheless worth exploring.

The Current Study

Given the challenges immigrants face in becoming part of Canadian society, both psychologically (in pursuit of community integration) and legally (in pursuit of national citizenship) at a time when their numbers are rapidly increasing, it is imperative to have a clear picture of how far native-born residents' acceptance extends, as this alone determines whether immigrants are seen as outsiders or equals. Furthermore, it is important to understand whether perceptions of immigrants' belonging are tied to their membership in a specific language group. This endeavor is particularly important in a time when citizenship is being increasingly reframed as a global rather than national concept (Rapoport, 2021). Therefore, the goals of this study were to explore how native-born residents of Canada use auditory and visual information (ethnicity and religion) to judge the citizenship of various immigrant L2 speakers, and to examine listener interactions with L2 speakers through the study of social networks, exploring relationships between those networks and citizenship ratings of specific L2 speaker groups.

One way to thoroughly investigate how biases against L2 speakers affect judgments of belonging in Canada is through investigations carried out in two representative environments. In consideration of the top immigrant host cities, Montreal and Calgary emerge as comparable East–West environments in terms of both general population (Montreal 1.8 million, Calgary 1.3 million) and immigrant population (Calgary 29%, Montreal 23%). What sets them apart are differences in sociolinguistic environment and overall diversity. For example, while Montreal is predominantly bilingual, Calgary is predominantly monolingual. However, Calgary has over 240 different ethnic communities, which is twice as many as Montreal (Statistics Canada, 2017). So, although Montrealers are more likely to speak an additional language, Calgarians have more opportunities for exposure to L2 speech, both of which have been identified as factors in listener

judgments (Yan & Ginther, 2017). What is clear across Canadian contexts, including the two chosen for this study, is that negative attitudes toward immigrants persist even in places of rich diversity (Kent, 2022), that they are often based on language, ethnicity, or religion (Esses, 2021; Godley, 2018; Nangia, 2013), and that they can be driven by everything from fears about limited resources like jobs (Danso & Grant, 2000; Montreuil et al., 2004), to fears about how an influx of new cultural practices will affect national identity (Louis et al., 2013).

To gain a comprehensive perspective of potential biases in Canada, residents of Calgary and Montreal (170 in each context) were presented with speech samples recorded by recent immigrants from four different linguistic backgrounds, representing either the top or fastest growing immigrant languages in common between the two cities (Arabic, Tagalog, Mandarin, and Punjabi), with North American English used for comparison. Participants were then asked to judge the speaker's citizenship status (alongside other dimensions not included in this study). Specifically, participants were asked to decide whether the speaker was a refugee, an immigrant, a temporary resident, a permanent resident, a naturalized Canadian citizen, or a Canadian-born citizen. This question was chosen because not only has citizenship been historically related to perceptions of group membership (Bloemraad, 2000) and national identity (Piller, 2001), but also because immigrants derive a sense of belonging from how they are perceived to fit into the host country's landscape (Simonsen, 2016). Sometimes the target speech samples were presented alone, and sometimes accompanied by images that matched the speaker's ethnicity, as in typical photo guise experiments (e.g., Kang & Yaw, 2021; McGowan, 2015), with or without an item of clothing (i.e., hijab, turban) that suggested religious affiliation. This design was implemented so that speech ratings could be explored in situations where the image prompted potential biases based on the speaker's ethnic and religious identity, relative to the no-image condition. The same

participants also completed a brief social network survey that explored their daily interaction with L2 speakers of English. Of key interest were rater judgments of various L2 speaker groups, their responses to the social network survey, and a potential relationship between these measures.

The following research questions were posed:

1. How do perceptions of L2 speakers' citizenship status compare for native-born residents of Montreal and Calgary when speech samples are presented in three conditions (audio-only, audio paired with a matching ethnic image, and audio paired with an image suggesting religious affiliation)?
2. How does interaction with L2 speakers (as shown through social network responses) compare for native-born residents of Montreal and Calgary, and how do those social network properties relate to perceptions of L2 speakers' citizenship status?

For the first research question, based on the difference in linguistic environments (predominantly monolingual vs. predominantly bilingual) between Calgary and Montreal, it was expected that Montrealers would be more generous in judgments of L2 speaker belonging. Native-born residents of the city, who recognize the value of linguistic diversity (Bousmah et al., 2021) and who possess multicultural identities themselves, might not only be less threatened by the arrival of immigrants, but also more accustomed to competing with other language groups for resources, such as jobs and political power (Bourhis et al., 2010). However, given the rich diversity in both cities and residents' regular exposure to minority ethnic and religious groups, any differences that surfaced were expected to be subtle. It was further expected that potential differences in citizenship evaluations for particular L2 speaker groups might be mediated by familiarity and presence of the L2 group in each city, such that Montrealers would elevate the

belonging of Arabic speakers and Calgarians would do the same for Tagalog, Punjabi, and Mandarin speakers (Statistics Canada, 2017).

Based on matched-guise research that has found evidence of negative bias toward L2 speakers when speech samples are accompanied by an image of a minority group member (e.g., Kang & Rubin, 2009; Kutlu, 2020), it was expected that judgments of belonging would differ for the same speech sample when presented as audio-only versus with an image. If so, it would be clear that appearance was swaying judgments held about speech alone (Lu & Gnevsheva, 2021). Furthermore, when speech samples were presented with an image of a person in a hijab or turban, it was expected that participants in both contexts—whose populations are both Christian majority (Statistics Canada, 2017)—would assign lower judgments of belonging to L2 speakers than when these speakers were heard in the audio-only or nonreligious conditions. Prior research has demonstrated that Christians often connect religion to national identity (Raney, 2009) and therefore share a belief that non-Christian immigrants threaten the stability of their nation’s established culture and traditions (McDaniel et al., 2011).

For the second research question, based on the rich ethnolinguistic diversity in both Montreal and Calgary, it was expected that the majority of residents in both cities would have at least some regular interaction with L2 speakers of English. It was further surmised, however, that the number of L2 social network contacts (or nodes) would be significantly higher for Montreal residents, as their regular contact would additionally include daily interaction with L2 speakers of English whose native language is French. As speakers of the official language of Quebec, Francophones represent 66% of Montreal’s population, making multilingual interactions highly likely. In contrast, 71% of Calgary’s population speaks English as a mother tongue, which suggests that monolingual English interactions might be most common. Interestingly, though,

considering Calgary's greater ethnolinguistic diversity, the contact that occurs could potentially involve a wider range of languages (Statistics Canada, 2017). Finally, although no prior research exists that explores the relationship between social networks of native-born residents in Canada and perceptions of belonging, it was expected that larger social network size and greater diversity would result in more favourable perceptions of belonging across L2 speaker groups as a byproduct of greater acceptance of immigrants overall. This idea is based on evidence in the fields of sociology and psychology showing the impact of diverse social networks on both inclusivity and reduced discrimination (Harell, 2010; Pauker et al., 2018; Vermeij et al., 2009).

Method

Raters

Raters included 340 native-born residents (170 Calgarians, 170 Montrealers) between the ages of 20 and 60. Altogether, there were 200 females, 131 males, and nine non-binary individuals. All used English as the primary language of interaction in their daily lives, though not all were native speakers of English (15 L2 speakers and 155 native speakers in each context). The choice to seek raters for whom English is the predominant language of communication, particularly in the multilingual setting of Montreal, was made to facilitate a clear comparison between contexts and to sidestep any English proficiency issues for L2-dominant raters, resulting in rater exclusion due to insufficient aural skills to judge L2 English speech. In addition, raters were born in the target environment and had been living there for at least three years prior to the study. Therefore, the rater group included lifetime residents as well as those who had spent periods away from the target context but returned, with an average length of residence of 25 years for Montrealers and 32 years for Calgary. Raters were recruited through social media, emails to local community organizations, word of mouth, and participant referrals. In both

contexts, raters were similar in their daily interaction with native English speakers, as reported on a 0–100 percentage scale ($M_{\text{Montreal}} = 86.00$, $SD = 17.99$; $M_{\text{Calgary}} = 82.93$, $SD = 16.86$).

However, raters differed in the number of languages spoken ($M_{\text{Montreal}} = 2.54$, $SD = 0.91$; $M_{\text{Calgary}} = 1.58$, $SD = 0.90$) and their L2 proficiency for those who identified as multilingual. This last measure was expressed on a 100-point scale, where 0 was labeled “can say a few words” and 100 was labeled “completely fluent” ($M_{\text{Montreal}} = 84.60$, $SD = 14.56$; $M_{\text{Calgary}} = 61.27$, $SD = 33.78$).

Materials

Speech Samples and Photo Guises

To collect speech samples suitable for the study, 10 Montrealers—one male and one female from each of the five target language backgrounds (Arabic, Mandarin, Punjabi, Tagalog, and North American English)—took part in an online mock job interview in English with an experienced interviewer. The L2 speakers, all born outside Canada, represented typical speech for the three most prominent and/or fastest growing immigrant language backgrounds in the two target contexts; native speakers of North American English provided a baseline comparison. All L2 speakers were noticeably accented, and they self-reported intermediate to advanced speaking proficiency. From each interview, a 30–45 second segment in which speakers discussed how they respond to constructive criticism was extracted. In exchange for their contribution, participants received both compensation and feedback on their interviewing skills.

Photos paired with the speech samples were chosen from a pre-rating task that included images from the Chicago Face Database (Ma et al., 2015), the Bogazici Face Database (Saribay et al., 2018), the London Face Database (DeBruine & Jones, 2017), and from images of several volunteers recruited by the researcher. All images showed people with minimal augmentations to appearance (e.g., plain hairstyles, no tattoos, minimal make-up), and individuals’ ages ranged

from 24 to 32 ($M = 27.08$, $SD = 2.40$). To create an even more unified set, a Photoshop professional altered each image to show each person against the same neutral background and in the same grey t-shirt (see Appendix A for examples). Thirty-eight images (at least 3 male and 3 female of each target ethnicity) were presented to 10 Montrealers between the ages of 24 and 55 ($M = 36.80$, $SD = 11.43$), who were asked to identify the person's ethnicity from among five broad groups (see Appendix B for the full list), then rate each photo for prototypicality, among other factors not considered in the final selection of images (e.g., status and solidarity traits). The final set of 10 images (1 male and 1 female for each of the five target ethnicities) was determined based on both high agreement across raters about the person's ethnicity as well as the highest rated ethnic prototypicality for a male and female within each ethnic group, with a score range of 65–84%. To incorporate suggestions of religious affiliation, Photoshop was used to add hijabs to females and turbans to males, for a final set of 20 images (10 nonreligious, 10 religious). Hijabs and turbans were selected specifically as religious symbols because they represent groups (Muslim women, Sikh men) that have been historically discriminated against in Canada (e.g., Aujla-Bhullar, 2020; Williams et al., 2022).

Speech Ratings

Raters experienced the audio narratives in three conditions (audio-only, audio with a matching ethnic nonreligious image, audio with a religious image), where each language group (Tagalog, Arabic, Punjabi, Mandarin, and North American English) was represented by one male and one female speaker. In the image conditions, each speech sample was paired with a photograph suggesting a potential visual depiction of the speaker, though none of the photos were of the actual speakers. To ensure that all possible pairings would be presented with the same frequency, 16 versions of the survey were created, and 10 raters were randomly assigned to

each version per location (Calgary, Montreal), for a total of 320 raters. This configuration resulted in a total of 80 target items (5 language groups \times 16 voice–image pairings), where the two speakers from each ethnic group were presented eight times with a nonreligious image and eight times with a religious image across all surveys. In addition, 80 other voice–image pairings included the same speakers presented with photos that did not match their ethnicity (with or without a religious head garment), but these mismatching conditions served as fillers and were not used for analysis. Because all possible combinations of voice–image pairings could not be presented in each survey (so that a given rater would not be exposed to the same speaker’s voice more than once), raters reacted to a random subset of five voice–image pairings, contributing 1,600 datapoints to the dataset. An additional 20 baseline raters (10 per city) completed an audio-only rating task in which all 10 speech samples (1 male and 1 female per language group) were presented in randomized order, contributing an additional 200 datapoints. Across all conditions, no rater heard the same audio or saw the same image more than once.

In response to each voice–image or audio-only presentation of the narratives, raters used a randomized multiple-choice list to judge the citizenship status of the L2 speaker from one of six options: refugee, immigrant, temporary resident, permanent resident, naturalized citizen, or Canadian citizen. This task was the last in a series of speech-based evaluations in which five 100-point scalar ratings were completed that addressed linguistic factors (comprehensibility, accentedness) and status and solidarity traits (intelligence, friendliness, trustworthiness), all of which were the focus of prior studies.

Social Network Survey

The social network survey (Appendix F), adapted for online use and inspired by Lybeck (2002) and Doucerain et al. (2015), explored raters’ contact with L2 English speakers. In

addition to reporting the number of L2 speakers with whom they interacted, at least on a weekly basis, raters also reported those speakers' native language(s), hours of interaction per week, and the language of interaction. Raters were also asked to report other information (reserved for another study) about their contact persons' ages and the role that each plays in their everyday lives (e.g., neighbour, co-worker, physician), as well as how close they are to that person on a scale of 1 (not very close) to 5 (very close). Finally, raters were asked broad, multiple-choice questions about which ethnic groups and religious backgrounds are represented across their networks.

Procedure

Raters completed all tasks through LimeSurvey in a session that lasted approximately 30 minutes and included a consent form, a background questionnaire, a social network survey, the speech rating task, and the social attitudes questionnaire (the data from which is discussed in another study). The choice to use strictly online tools was born part out of practicality and part out of necessity due to COVID-19, which limited opportunities for in-person interaction at the time of data collection. In general, online tools have shown high internal consistency comparable to lab-collected data (Gosling et al., 2004; Nagle, 2019). For added security, however, several additional controls were added. Raters were not allowed to return to completed sections, change their answers, or skip audios or scales; their progress was time-tracked, they were asked to use personal headsets for listening or to complete the task in a quiet location, and they were encouraged to report any noise or interruptions that occurred during the rating session.

Data Analysis

Evaluations of L2 speaker citizenship were converted to numerical rankings, with “refugee” at one end of the continuum (1) followed by “immigrant” (2), “temporary resident”

(3), “permanent resident” (4), “naturalized citizen” (5), and “Canadian-born citizen” (6). Social network survey responses were analyzed separately per context to derive six measures that were expected to be most relevant to evaluations of citizenship status: (a) total number of nodes (contacts) in each rater’s network, (b) total number of language varieties in the network, (c) language diversity per network, computed as a type–token ratio of language varieties over the total number contacts reported, (d) total number of hours each rater interacted with L2 speakers per week, (e) total number of religious groups represented across each rater’s network, and (f) religious diversity per network, also computed as a type–token ratio of religious varieties over the total number contacts reported.

To address the first research question, data analyses were carried out through mixed-effects modeling using the lme4 package (Bates et al., 2014) in R version 4.0.2 (R Core Team, 2020). Although there is no consensus regarding the criteria for considering statistical significance in mixed-effects modeling, with some scholars using $t > |2.00|$ as the benchmark of significance (Linck & Cunnings, 2015), we examined 95% confidence intervals (CIs) to check the statistical significance of each parameter (interval does not cross zero). To account for random variance, all models included random intercepts for participants (i.e., raters) and, where applicable, for items (i.e., speakers). Random slope models were also examined; however, the inclusion of random slopes did not improve model fit, so the final models excluded random slopes. To address the second research question, social network measures were first analyzed descriptively for between-context comparison, then Pearson correlations were carried out to examine potential relationships between raters’ social network characteristics and their perceptions of L2 speaker citizenship.

Results

Citizenship Labels

The first research question explored how native-born Canadian residents evaluate L2 speakers' citizenship status for various ethnolinguistic groups when their speech is presented in three visual conditions. The citizenship labels (summarized in Table 1) were first analyzed separately within each speaker group, with condition (audio-only, matching ethnic image, religious image) and location (Calgary, Montreal) used as fixed effects and an intercept for rater and an intercept for speaker included as random effects. Ultimately, while none of the speaker groups were perceived to be refugees, all were perceived differently in terms of citizenship, with English speakers rated highest in terms of belonging and Punjabi speakers rated lowest.

Table 1

Means (Standard Deviations) for Citizenship Labels by Speaker Group, Condition, and Location

Speakers	Calgary			Montreal		
	Audio-only	Nonreligious	Religious	Audio-only	Nonreligious	Religious
Arabic	2.90 (1.33)	2.84 (1.25)	2.96 (1.41)	2.50 (0.95)	2.88 (1.23)	3.22 (1.37)
Mandarin	3.60 (1.57)	3.60 (1.45)	3.49 (1.58)	4.00 (1.62)	3.25 (1.33)	3.45 (1.33)
English	5.45 (1.05)	5.55 (0.97)	5.75 (0.68)	5.85 (0.49)	5.78 (0.57)	5.60 (0.85)
Punjabi	2.45 (1.19)	2.67 (1.10)	2.74 (1.26)	2.50 (1.15)	2.80 (1.22)	2.58 (1.04)
Tagalog	4.45 (1.47)	4.36 (1.56)	4.26 (1.58)	4.65 (1.42)	4.45 (1.65)	4.42 (1.57)

As summarized in Table 2, there was no significant effect of condition for any speaker group, such that the label choices in the nonreligious and religious image conditions were comparable to those in the audio-only condition. In terms of location, Calgary and Montreal

raters also provided similar label choices for all speaker groups, regardless of the condition (for summary of full mixed-effects models, see Appendix G).

Table 2

Summary of Mixed-Effects Models for Citizenship Labels in the Nonreligious and Religious Conditions by Location and Speaker Group

Speakers	Nonreligious			Religious			Location		
	<i>Estimate</i>	<i>t</i>	<i>p</i>	<i>Estimate</i>	<i>t</i>	<i>p</i>	<i>Estimate</i>	<i>t</i>	<i>p</i>
Arabic	0.16	0.71	.481	0.39	1.78	.077	0.09	0.67	.502
Mandarin	-0.38	-1.24	.216	-0.33	-1.10	.274	-0.15	-0.97	.333
English	0.01	0.09	.928	0.03	0.18	.857	0.08	0.94	.349
Tagalog	-0.14	-0.65	.515	-0.21	-0.94	.350	0.13	1.01	.311
Punjabi	0.26	1.21	.232	0.18	0.83	.408	-0.01	-0.10	.922

Note. The baseline (audio-only) condition and the Calgary location were set as reference levels. Full statistical output appears in Appendix G.

Because raters chose similar citizenship labels regardless of the image condition in which the speech was presented, condition was not considered in further analyses of citizenship evaluations. Instead, subsequent modeling explored whether raters differed in their evaluations of the five speaker groups. In these analyses, speaker group (Arabic, Mandarin, Punjabi, Tagalog, North American English) and location (Calgary, Montreal) served as fixed effects, and an intercept for rater and an intercept for speaker were again entered as random effects (see Appendix G for full model output). While location did not emerge as a significant predictor of

citizenship labeling (Estimate = 0.03, 95% CI = [-0.11, 0.17], $t = 0.397$, $p = .692$), speaker group did (Estimate > |0.27|, $p < .002$).

To understand relative differences in rater evaluations of the speaker groups, between-group contrasts were explored through analyses of estimated marginal mean differences (EMM_{diff}) using the Satterthwaite method with Tukey adjustment for multiple comparisons (see Appendix G for full summary). In terms of citizenship labels, both Montreal and Calgary listeners (i.e., no difference by context) gave significantly different labels to all speaker groups. The English speakers were considered more likely to be Canadian-born citizens, with a mean rating close to 6.00 ($M = 5.67$, $EMM_{diff} > 1.27$, 95% CI = [1.01, 3.26]), the Tagalog speakers were labeled as permanent residents, with a mean rating close to 4.00 ($M = 4.39$, $EMM_{diff} > 0.91$, 95% CI = [0.65, 1.98]), the Mandarin speakers were categorized midway between temporary and permanent residents ($M = 3.49$, $EMM_{diff} > 0.54$, 95% CI = [0.28, 2.44]), the Arabic speakers were labeled as temporary residents, with a mean rating close to 3.00 ($M = 2.94$, $EMM_{diff} > 0.27$, 95% CI = [0.01, 2.98]), and the Punjabi speakers elicited the description that fell between immigrants and temporary residents ($M = 2.67$, $EMM_{diff} > 0.27$, 95% CI = [0.01, 3.26]). Thus, for citizenship labels, the ranking was English (Canadian-born citizens) > Tagalog (permanent residents) > Mandarin (permanent/temporary residents) > Arabic (temporary residents) > Punjabi (temporary residents/immigrants).

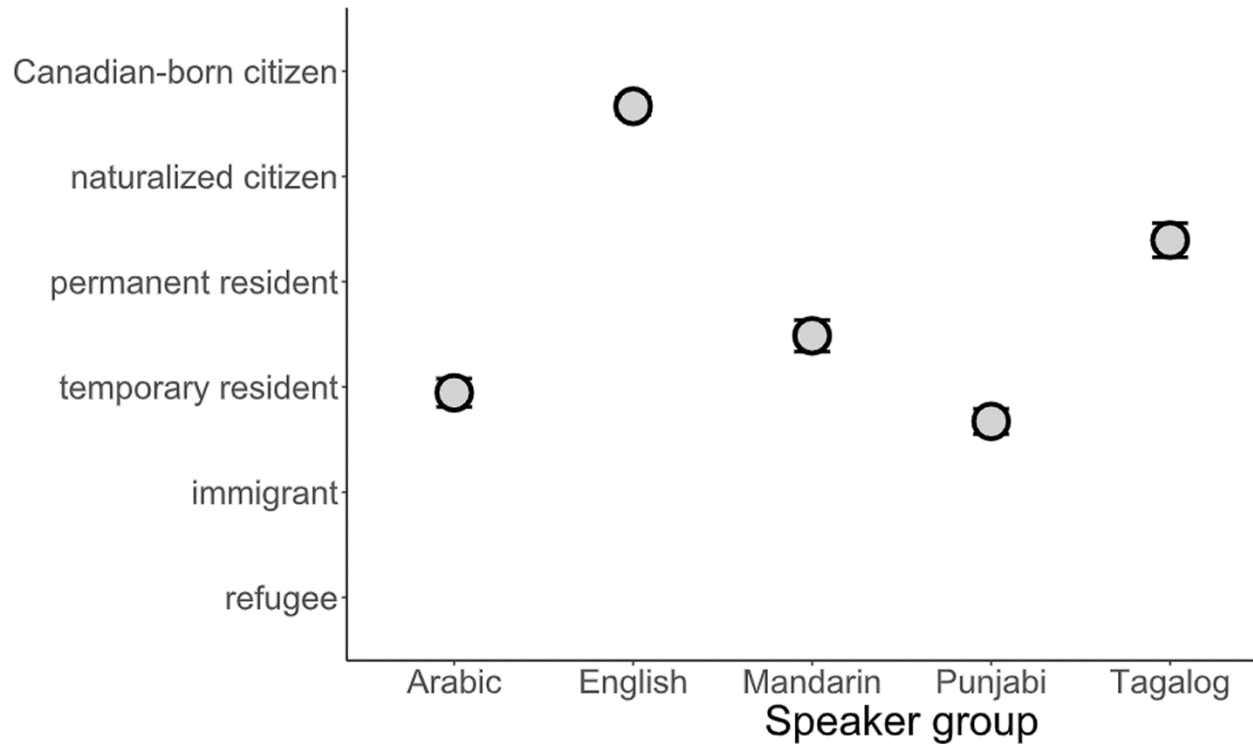


Figure 1

Citizenship Labeling by Speaker Group, with Brackets Enclosing 95% Confidence Intervals

Social Network Surveys

The second research question asked about raters' interactions with L2 speakers in two Canadian contexts and explored potential relationships between those social network characteristics and their perceptions of L2 speakers' citizenship. Responses to the social network survey (summarized in Table 3) revealed small but significant differences between the two contexts, showing that while Montrealers reported more L2 contacts and more hours of interaction overall (although there was a wide range of contact hours in both contexts), Calgarians reported interactions with speakers of a wider variety of languages, which suggests that the average L2 social network in Calgary is more diverse. Furthermore, although

Montrealers reported more religious variety among their L2 contacts, there was slightly more religious diversity in Calgarians' L2 networks.

Table 3

Means (Standard Deviations) and Between-Context Comparisons for Raters' Responses to the Social Network Survey

Variable	Location		Comparison		
	Calgary	Montreal	<i>t</i>	<i>p</i>	<i>d</i>
Nodes	3.42 (3.43)	4.32 (3.90)	-23.71	< .001	-0.25
Language variety	2.29 (2.11)	2.05 (1.70)	12.49	< .001	0.13
Language diversity	0.64 (0.36)	0.52 (0.33)	33.13	< .001	0.34
Hours per week	29.73 (49.61)	39.15 (62.97)	-16.07	< .001	-0.17
Religious variety	1.81 (1.69)	2.04 (1.68)	-11.02	< .001	-0.11
Religious diversity	0.56 (0.38)	0.53 (0.34)	7.32	< .001	0.08

Social Network Characteristics and Relationship to Citizenship Labeling

Pearson correlations (summarized in Table 4) were then used to explore potential relationships between raters' responses to the social network survey and their evaluations of speaker citizenship status (regardless of the visual condition). None of the associations reached the .25 benchmark for a weak effect (Plonsky & Oswald, 2014), suggesting that no social network measure had a meaningful association with raters' labeling behaviours.

Table 4

Pearson Correlations Between Raters' Citizenship Evaluations and Responses to the Social Network Survey

Variable	Arabic	Mandarin	English	Punjabi	Tagalog
Calgary					
Nodes	-0.10	0.13	-0.05	0.00	0.13
Language variety	-0.04	0.07	-0.02	-0.01	0.09
Language diversity	0.09	-0.02	0.11	0.09	0.08
Hours per week	-0.01	0.12	-0.13	-0.08	-0.03
Religious variety	-0.05	0.06	0.02	-0.05	0.02
Religious diversity	0.08	0.02	0.01	0.00	0.02
Montreal					
Nodes	0.12	0.01	0.03	0.03	0.01
Language variety	0.02	-0.08	0.15	-0.01	-0.13
Language diversity	-0.08	-0.08	0.12	0.01	-0.16
Hours per week	0.12	-0.01	-0.06	0.16	0.03
Religious variety	0.03	-0.07	0.18	0.10	-0.07
Religious diversity	-0.09	-0.17	0.16	0.11	-0.07

Discussion

This study's goal was to explore how English-dominant native-born residents in two Canadian contexts compare in judgements of L2 speaker citizenship as well as how those judgments might be related to raters' L2 social networks. There were no significant differences between speech samples presented as audio-only versus with a nonreligious or religious image of a person that matched the speaker's ethnicity. There were also no differences between Calgaryans

and Montrealers in their judgments of L2 speaker citizenship. However, notable differences emerged across raters in how the citizenship of specific language groups is viewed. In addition, significant differences between contexts surfaced regarding native-born residents' interactions with L2 speakers, insofar as Montreal raters had more contacts and more hours of interaction with L2 speakers as a whole but Calgary raters' networks were proportionally more ethnically and religiously diverse. Finally, there was no relationship between social network characteristics and citizenship ratings in either context.

Citizenship Labels

Of primary interest for our first research question was whether raters in Montreal and Calgary would differ in their judgments of L2 speaker citizenship and whether ratings in either context would be influenced by the presentation of images (nonreligious or religious) alongside speech samples. Ultimately, there was no image effect. Citizenship labels assigned to each speaker were not influenced by whether the speech sample was presented alone, with an image that matched the speaker's ethnicity, or with an image that suggested the speaker's religious affiliation. This finding, although consistent with the ratings of the same speakers' linguistic and personal characteristics provided in the same sessions (and reported in separate manuscripts), is incompatible with prior research that has found an image effect on speech ratings (e.g., Hansen et al., 2017b; Ghanem & Kang, 2021; Hu & Lindemann, 2009; Rubin, 1992; Rubin & Smith, 1990). For example, Kutlu (2020) invited 67 native English speakers to judge the accentedness of American-English and Indian-English speakers when presented with either a White or a South Asian face. The speech samples were rated as more accented when presented with a South Asian face, and this effect was particularly prominent for the American-English speech samples.

There are several explanations for why the present findings may have diverged from those reported previously. First, most of the studies that have demonstrated an image effect on speech ratings (including those mentioned above) were carried out in the United States where residents—even those of diverse racial and ethnic backgrounds—typically think of someone who is White when asked to describe an American (Devos et al., 2010; Devos & Banaji, 2005). This outlook suggests that appearance plays a prominent role in the perception of US citizenship. While some research in Canada has also revealed appearance-based biases in regard to healthcare access (e.g., Veenstra, 2011), the nation has worked hard to infuse its identity with multiculturalism (Brosseau & Dewing, 2018) and inclusion (Government of Canada, 2018). It is therefore possible that a person’s appearance is not at the top of the list when Canadians assess citizenship.

Second, it is possible that rater biases might not have been fully activated if raters were not certain about the ethnic group to which each image belonged. Although all images were assessed through pre-rating as moderately to highly prototypical of the target ethnic group, there was a wide range of prototypicality scores (65–84%), so some depicted individuals may not have been perceived as belonging to a particular target ethnicity. Third, it is possible that given the two sources of input (audio and visual), raters simply believed speech to be a better determiner of someone’s citizenship. Indeed, speech can offer many clues about a person’s group membership, including their ethnicity and country of origin (Risager, 2006), and prior research has documented that listeners sometimes ignore visual information when auditory information is present (McGowan, 2015; Rakić et al., 2011). In other words, while someone might *look* foreign, that does not mean they will *sound* foreign (Uther et al., 2006), and raters might rely on voice as the more reliable measure of group membership.

It is interesting that visible religious symbols, such as hijab and turban, also had no influence on citizenship ratings. This outcome might suggest that although religious discrimination continues to surface in Canadian society (e.g., Rahmath et al., 2016; Vickers, 2004), Canadian residents might be particularly sympathetic at a time when government policies, such as Quebec's Bill 21, openly target visible religious minorities by preventing public workers like teachers, healthcare professionals, and police officers from wearing anything indicative of their religious affiliation. Finally, although speech samples were played only once, there was no time limit for raters to respond, giving them plenty of time to consider the relative weight of each input factor: voice, ethnic appearance, and (in half of the samples) religious affiliation. Without time pressure, which encourages a less consciously controlled response (Levinson et al., 2019), raters may have deliberately avoided any suggestion of bias based on a speaker's appearance and religious beliefs.

In between-context comparisons, results also showed that Montrealers and Calgarians assigned similar citizenship labels to each L2 speaker. This is not altogether surprising, as residents of both cities have regular exposure to all of the L2 speaker groups included in this study, albeit in different proportions (Statistics Canada, 2017). The differentiating factor that prompted this comparison, however, was sociolinguistic context. It was unknown whether native-born residents in a predominantly monolingual environment (Calgary) versus a predominantly bilingual environment (Montreal) would differ in the way they viewed L2 speaker citizenship. Here, the linguistic environment does not appear to have played a role. For one, although being multilingual can temper one's acceptance of others (Mepham & Martinovic, 2018), appreciation for multiple cultures is, of course, not exclusive to those who speak multiple languages. Acceptance can also spring from such factors as intergroup contact (Al Ramiah &

Hewstone, 2013) and education (Verkuyten & Thijs, 2013). Given the two cities' comparable immigrant populations, historical presence of immigrants, and prioritization of welcoming immigrant policies (Guo & Guo, 2016; Viola, 2012), it makes sense that residents would share common perspectives about citizenship. Furthermore, both cities are progressive forces in their more conservative provinces, boast high cultural diversity, and benefit from immigrant contributions to the economy (Statistics Canada, 2017). So, although acts of discrimination and barriers to immigrant success are common in both contexts, the two host communities have also demonstrated acceptance of newcomers that is in line with Canada's mission as a whole (Brosseau & Dewing, 2018).

Language-Based Citizenship Hierarchy

Although raters in both contexts were similar in their citizenship judgments of the same speakers experienced across different audio and visual configurations, they collectively viewed the five speaker groups as having distinct levels of citizenship. They assigned the highest level of belonging (Canadian-born citizen) to the native English speakers, followed by the Tagalog speakers (permanent resident), the Mandarin speakers (permanent/temporary resident), the Arabic speakers (temporary resident), and finally, the Punjabi speakers (temporary residents/immigrants). None of the L2 speaker groups were judged to be refugees. To explain this citizenship hierarchy, we might first consider familiarity. After all, native English speakers, rated highest in terms of citizenship, share in-group membership with the raters. However, Tagalog speakers, judged to be closely behind English speakers in terms of citizenship, represent the seventh most spoken immigrant language in Canada, whereas Punjabi speakers, rated the lowest in terms of belonging, represent the second, although the language is more common in Calgary than in Montreal (Statistics Canada, 2016). Since rater familiarity with each language

group does not appear to have been a factor, it is perhaps wiser to examine what fuels perceptions of citizenship.

First, there is disparity between documented criteria for national citizenship and criteria that an individual community (e.g., nation, province, city) considers important to citizenship (Simonsen, 2016). For instance, while Canada prioritizes residency, knowledge of the country's laws and history, and both occupational and language skills (Government of Canada, 2022), Canadian citizens connect citizenship to national identity, which is increasingly defined by narrow criteria, such as birthplace and religion (Raney, 2009). According to social identity theory (Tajfel 1959, 1972), group identity is derived from commonalities, which might explain why, given Canada's Christian-majority population, with 55% in Calgary and 66% in Montreal (Statistics Canada, 2017), Tagalog speakers, many of whom speak English from birth and as many as 97% of whom are Christians, would be perceived as closer to Canadian-born citizens than Punjabi speakers, the majority of whom are Sikh, Hindu, or Muslim and work hard to preserve ties to their motherland (Lindsay, 2007). In other words, while raters might not have demonstrated explicit religious preference by allowing images of a person in a hijab or turban to influence their judgments, the sound of a Punjabi speaker might have triggered an implicit association with a religious and/or ethnolinguistic out-group that resulted in perceptions of limited belonging.

While religious group membership might also partially explain why Arabic speakers were labeled as temporary residents—although to a lesser degree, as 44% of Arab-origin Canadian residents are Christian—we might also look to group threat theory (Blalock, 1956; Blumer, 1958). In liberal democratic countries like Canada and the United States, perceived threat is used as justification for exclusion of immigrants from acceptance and, most certainly,

citizenship (Raijman et al., 2008). Often, group threats are fueled by concerns about national security and political power (Albarello & Rubini, 2018). Since the September 11, 2001 attacks in the United States, many Canadians have associated people of Arab origin, even those who are Canadian-born, with terrorism, and therefore perceive an entire ethnolinguistic group to be a threat to Canada's safety (Razack, 2007). So, even though Canadians of Arab origin are highly educated, make substantial contributions to Canada's economy, and are politically and civically active (Lindsay, 2007), perception of threat might restrict native-born residents' judgments of their belonging. From the perspective of group threat theory, Punjabi speakers might also be assigned lower status of belonging, as this immigrant group is becoming increasingly politically active in Canada. In fact, Punjabi is the third most spoken language in Parliament (Singer, 2019). So, while Punjabi speakers contribute substantially to Canadian society, it is possible that their political power as members of a minority group is a threat sufficient enough for some Canadians to assign judgments of limited belonging.

Finally, Mandarin speakers were also assigned a label suggesting less permanence (temporary/permanent resident) than native English speakers or Tagalog speakers, which is surprising considering that Canada has a long history of Chinese immigration, and people of Chinese origin represent the largest non-European immigrant group in Canada. However, the Chinese population is growing significantly faster than the overall Canadian population (Lindsay, 2007), which could be perceived as threat to national identity due to sheer numbers. Added to this is the Chinese community's tendency toward segregation, along with a uniquely Chinese identity linked to home language use and preference for Chinese media outlets, which inevitably results in weakened ties with Canadian society (Deng & Walker, 2007; Derwing & Waugh, 2012). In addition, Chinese immigrants to Canada often maintain lower English

proficiency than other L2 speaker groups (Derwing et al., 2010), which arguably limits interaction with native-born residents and potentially influences perceptions of belonging. In terms of government requirements, limited language proficiency and limited knowledge of Canadian society most certainly hinder citizenship attainment. Collectively, these factors might explain why Chinese immigrants have experienced the sharpest decline in Canadian citizenship rates of any immigrant group in recent years, which Hou and Picot (2019) attribute, among other factors, to the lure of China's strengthening economy and immigrants' lower proficiency in either of Canada's two official languages. So, while the number of Chinese immigrants in Canada is rapidly increasing, their status as permanent members of Canadian society is not.

Social Networks

The goal of the second research question was to compare social networks of residents in Montreal and Calgary in terms of regular interaction with L2 speakers. Here there were significant differences between contexts. Montrealers averaged more L2 contacts (4.32 vs. 3.42) and more hours of interaction per week with these contacts (39.15 vs. 29.73), but Calgarians showed more linguistic and religious diversity among their contacts. Overall, this outcome reflects the large immigrant populations in both cities but aligns with the existence of more language communities in Calgary (240 vs. 120). In fact, Calgary's diversity might have offset the benefits of multilingualism in Montreal to result in similar citizenship ratings of L2 speakers, as both living in a diverse environment and being multilingual have been associated with greater acceptance of out-group members (Mephram & Martinovic, 2018; Al Ramiah & Hewstone, 2013), although this was not directly investigated.

An additional objective was to explore possible associations between native-born residents' social network characteristics and their choice of citizenship labels for L2 speakers.

No significant correlations surfaced in either Montreal or Calgary. This might best be explained by the idea that multicultural interaction, practiced by many Canadians and encouraged by Canadian policy (Berry, 2013), is not the same as sharing all rights and resources that accompany Canadian citizenship. In other words, while native-born residents happily share physical space, especially given the many contributions immigrants make to Canadian society, this might not correspond with a desire for full equality and shared national identity (Raijman, 2008), a stance that would require native-born residents to relinquish social dominance (Esses et al., 2006). Therefore, it is understandable that how residents interact with immigrants might not correlate with their perceptions of immigrant belonging.

Implications

By the year 2040, Canada's population growth is expected to be driven entirely by immigration (Vomiero, 2018). With Canada's future so heavily dependent on immigrant success, it is vital to not only have a highly nuanced understanding of how to forge effective relationships between native residents and newcomers, but also to clear the obstacles that so often block immigrants' paths. Continued education, awareness raising, and activities that encourage positive interaction (like shared community projects) are therefore essential for building positive, successful connections between native-born residents and newcomers. These cross-cultural opportunities are especially important for ethnic groups that do not commonly interact, because positive interaction can remove fears that are based on stereotypes instead of experience and turn perceived threat into an opportunity to learn (Bochner, 1982). In addition, given that Canadian immigrants report that most discrimination occurs in the workplace (Lindsay, 2007), employers bear the responsibility of decreasing negative attitudes through documented anti-discrimination policies, guidelines for resolution, and continued education, to include acceptance of increasingly

common nonstandard varieties of English (Suraweera, 2015). These actions might be most effective as part of a more substantial journey toward a reimagined national in-group, to include immigrants, which could boost reciprocal perceptions of belonging between native-born residents and immigrants (Esses et al., 2006).

Perhaps most troubling about the citizenship labeling of particular L2 speaker groups is that Canadians whose ethnic origin is represented by these languages report a strong sense of belonging to Canada, in most cases substantially stronger than to their own ethnic or cultural group (Lindsay, 2007). Thus, there is incongruity in the way native-born residents see immigrant belonging and the way immigrants see themselves, which could offer newcomers a false sense of security about how they fit into the fabric of Canada's identity. This disparity also suggests that immigrants' allegiance and contributions might be substantially underappreciated by the host community, particularly considering that the vast majority of immigrants (86.2%) do, in fact, become Canadian citizens (Hou & Picot, 2019).

Limitations and Future Research

There are a few important limitations of this study worth noting. First, the two contexts chosen, though different linguistically, are richly diverse with long histories of welcoming immigrants, which likely resulted in similar views about Canadian citizenship. Future research should explore these perceptions in more contrastive environments, such as rural versus metropolitan areas or involving cities with fewer immigrants, monolingual environments, or both. While this study has furthered our understanding of Canadian perspectives on citizenship as a whole, it remains important to fully explore how different environments can contribute to perceptions of belonging, which should include thorough investigation of the role of exposure to L2 speakers through mapping of individual or group social networks. Second, there was only one

male and one female speaker per language group, which prevented any gender-based comparisons. It also meant that feelings about an entire language group's citizenship were based on the voices of two speakers. Future research should explore native-born and immigrant participants' perceptions of citizenship status with a much larger group of representative L2 speakers. Third, the role of language proficiency in a speaker's citizenship status was not explored here, but it would be important to do so in order to fully understand how such an important factor in the granting of legal citizenship influences perceptions of belonging held by the host community.

In addition, the multiple-choice options presented as citizenship labels were not defined, such that, for instance, there was no explanation of the nuances between temporary and permanent resident. In the absence of definitions, participants needed to rely on their own perception of the differences between categories, which presumably reveals how native-born residents view belonging. Of particular concern here is that participants varied in what different labels meant for them and how those labels were different from each other. Although capturing participants' impressions of each label provides us with a snapshot of their understanding without further information, it would also be important to explore how native-born residents respond when they are provided with the legal definitions of such terms. Finally, despite encouraging results regarding the effect of ethnic and/or religious appearance on citizenship status of L2 speakers, these judgments were made in an unnatural environment where the participants had unlimited time to think about their ratings and where speakers were evaluated alongside filler items illustrating less common ethnic and religious image combinations (e.g., Chinese man in a turban), so judgments may or may not fully reflect participants' real-world behaviour. Given the high stakes of such research, it would be important to continue to explore

judgments of citizenship and belonging through multiple methods including, for instance, implicit association tasks or judgments completed under time pressure.

Conclusion

The current study investigated native-born residents' judgments of L2 speaker citizenship and explored social network interaction in two Canadian contexts. In judging citizenship, raters were not influenced when the speech sample was accompanied by an image matching the speaker's ethnicity or showing religious affiliation. There was also no difference between contexts, in the sense that Montreal and Calgary raters were in agreement about the citizenship ratings. However, significant differences emerged regarding the citizenship labeling of specific L2 speaker groups, with native English speakers rated as Canadian-born citizens and Punjabi speakers labeled between temporary residents and immigrants—the statuses furthest from Canadian citizenship. There were also significant differences in social network characteristics between Montreal and Calgary, with Montrealers having larger L2 speaker networks and more contact overall, but Calgarians having networks that were more linguistically and religiously diverse. Ultimately, though, there were no associations between raters' social networks and citizenship ratings. These results further our understanding of the role of speech, ethnic appearance, and religious affiliation in judgments of various L2 speaker groups and, although intergroup contact is plentiful in both cities, not all L2 speaker groups are viewed the same way in terms of belonging.

CHAPTER 5

General Discussion and Conclusion

Overview of Studies

The overall goal of this dissertation was to explore social attitudes held by native-born residents of two Canadian contexts, Calgary and Montreal, as representatives of their individual cities and of Canada as a nation. Of particular interest was how residents judge the speech of various L2 groups in terms of linguistic factors (comprehensibility, accentedness), status and solidarity traits (intelligence, friendliness, and trustworthiness), and citizenship status, as well as how residents interact with L2 speakers and what their social attitudes are toward immigrants overall.

Key Findings

Collectively, the three studies in this dissertation contribute to our understanding of the roles of sociolinguistic context and visual information in L2 speech ratings, the attitudes toward particular L2 speaker groups, and the influence of attitudes and social network exposure on judgments of L2 speech. The first key finding was that residents of Calgary and Montreal were generally in alignment in their opinions about the speaker groups. They judged the speakers similarly across rated dimensions (comprehensibility, accentedness, intelligence, friendliness, trustworthiness, citizenship), with two subtle differences, both reported in Study 1, where Montreal raters found the Arabic speakers to be more comprehensible and the Mandarin speakers to be more accented than Calgary raters did, regardless of image condition. In the interest of exploring a broader Canadian perspective in Study 2, all data were pooled from both locations for status and solidarity ratings; therefore, contextual comparisons were not investigated. The second key finding was that there was no significant effect of visual information on ratings of L2

speech across conditions, which were defined through images matching or mismatching the speaker's ethnicity in Study 1 and through images that were nonreligious (but matching in ethnicity) or religious in Studies 2 and 3.

Upon finding no significant effect of context or condition on ratings and therefore pooling the data, a third key finding surfaced regarding differences in how raters judged each speaker group. The native speakers of North American English (baseline group) were rated the most comprehensible and the least accented (Study 1). They also shared the most favourable ratings for intelligence and friendliness with the Arabic and the Tagalog speakers, as well as with the Tagalog speakers for trustworthiness (Study 2). In terms of citizenship, they were labeled to most likely be Canadian-born citizens (Study 3). The Tagalog speakers were rated the second most favourably in terms of comprehensibility and were judged to be only slightly more accented than the English speakers. They also shared the highest ratings of intelligence and friendliness with the English and the Arabic speakers, and with the English speakers for trustworthiness. They were judged only slightly less favourably than the English speakers in terms of citizenship, with raters believing that the Tagalog speakers were most likely permanent residents.

The Arabic and Punjabi speakers were judged least favourably in terms of comprehensibility and were judged to be most accented, with the Punjabi speakers rated as slightly more accented than the Arabic speakers. However, ratings of these two speaker groups diverged for status and solidarity. While the Arabic speakers shared the most favourable ratings of intelligence and friendliness with the English and Tagalog speakers and were rated only slightly less favourably for trustworthiness, Punjabi speakers shared the lowest intelligence and trustworthiness ratings with the Mandarin speakers and were rated only slightly higher than the Mandarin speakers in terms of friendliness. The Arabic and Punjabi speakers were rated

similarly in terms of citizenship, however, with raters believing that the Arabic speakers were most likely temporary residents and that the Punjabi speakers were somewhere between immigrants and temporary residents. Finally, while the Mandarin speakers fell largely in the middle of the rating scales for comprehensibility, accentedness, and citizenship (judged to be between temporary and permanent residents), they were rated the lowest in terms of friendliness and shared the lowest ratings for intelligence and trustworthiness.

The fourth key finding is that residents of Montreal and Calgary were similar in their social attitudes toward immigrants, with a few subtle differences. In Study 1, raters generally agreed that individual characteristics and contributions of immigrants were moderately important, that their race and religion were relatively unimportant, and that they generally do not mind when immigrants occupy prominent roles in society. Compared to Calgarians, Montrealers placed greater importance on desired immigrant characteristics and immigrants' race and religion and expressed marginally stronger acceptance of immigrants in prominent societal roles. Study 2 revisited the societal roles category—showing high acceptance of immigrants in prominent societal roles across rater responses—for the purpose of comparison to status and solidarity ratings. Two additional social attitudes categories were also investigated in Study 2, where rater responses revealed moderately high perceptions of immigrants' civic contributions to Canada (i.e., overall quality of life, contributions to society, and economic impact) and placed moderately high importance on immigrants having political involvement in Canada (i.e., becoming citizens and voting in elections).

In Study 3, the L2 social networks of raters were examined, revealing small but significant differences between the two contexts, showing that while Montrealers reported more L2 contacts and more hours of interaction overall (although there was a wide range of contact

hours in both contexts), Calgarians reported interactions with speakers of a wider variety of languages, which suggests that the average L2 social network in Calgary is more diverse. Furthermore, although Montrealers reported more religions represented overall among their L2 contacts, there was slightly more religious diversity in Calgarians' L2 networks. Across studies, however, no significant associations emerged between social attitude responses or social network characteristics and ratings of L2 speech.

Overall Implications

Understanding Attitudes in Canada

Results across the three studies have shown that there is considerable common ground between Montreal and Calgary, reflective of Canada's overall modern identity as a multilingual, multicultural nation accepting of newcomers (Lund & Hira-Friesen, 2014). This alignment suggests that perhaps efforts made at the community, provincial, and national level toward inclusivity are working, at least in regions that host large immigrant populations. Furthermore, these findings support the probability that multiple factors (e.g., education, interaction) are at work in diverse communities to offset any differences that might surface based on sociolinguistic context. Both predominantly bilingual Montreal and predominantly monolingual Calgary expressed attitudes toward L2 speakers across all measures that were not only similar, but generally positive. These findings also provide support for the importance of establishing regional and national identities that celebrate diversity, particularly in parts of the world where regular interaction with minority groups is limited.

Although residents of the two contexts highlighted in this dissertation were similar in their judgments, an important finding surfaced regarding the rankings of various speaker groups, with a clear distinction between preferred (or dispreferred) varieties of English that held across ratings of language characteristics, status and solidarity traits, and citizenship status. This

supports prior research that has not only revealed accent as a conspicuous signal of group membership (Kang et al., 2010) but has also established the key role of language as a categorization factor (e.g., Beaulieu et al., 2022; Dragojevic & Goatley-Soan, 2020). For example, Klitmøller et al. (2015) investigated language-based social categorization in a Finnish multinational corporation and found that employees from five different European countries expressed low tolerance for language proficiency issues and made categorical judgments about the competence of particular language groups based on their telephone interactions with individual speakers, despite having regular interactions with co-workers in English. In addition to language's role in categorization, it is possible that the way particular ethnic groups interact in Canadian society and align with national identity informs native-born residents' judgments as triggered by the sound of an individual speaker. For example, the L2 speaker groups focused on in this dissertation represent various levels of educational attainment, participation in Canada's civic and political endeavors, and alignment with Christian-majority religious beliefs (Lindsay, 2007), all of which might affect how they are ranked across multiple social categories based on impressions native-born residents have formed about members of those L2 speaker groups as a whole.

Of further interest is the robust finding across studies that inclusion of an image did not affect rater judgments of linguistic or extralinguistic factors. It made no difference whether a speech sample was presented as audio-only, with a matching or mismatching ethnic image, or with an image suggesting religious affiliation. Although this finding runs counter to many studies involving audiovisual stimuli in which speakers were upgraded or downgraded when an image was present (e.g., Kutlu, 2020; Lu & Gnevsheva, 2021; Lybaert et al., 2022; Rubin & Smith, 1990), it aligns with prior research that has shown speech to be a primary cue to categorization,

more salient to a listener than other forms of information or stimuli (e.g., Hansen et al., 2017b; Rakic et al., 2011). Furthermore, it is worth considering that Canadian raters, immersed in a largely welcoming environment for newcomers (Guo & Guo, 2016; Scott, 2020), may behave differently than raters in other contexts. For example, many studies showing an image effect have been conducted in the United States, a country with a decidedly more conservative outlook on diversity (Harell et al., 2012).

Theoretical Implications

In consideration of the inevitable competition for resources such as jobs and housing and the potential shift in societal identity that occurs as a population grows and changes, as with a rapidly growing immigrant population, both group conflict theory (Sherif, 1966) and group threat theory (Blalock, 1956; Blumer, 1958), which have been used to explain biases held by majority group members toward members of various out-groups, were outlined as frameworks for this research. Given various tensions that have been documented to exist in Canada between language groups (Swain & Lapkin, 2005), ethnic groups (Kymlicka, 2009), and religious groups (Vang et al., 2019), it was expected that group biases might surface in ratings of L2 speech as a result of the image guise. If they had, that would have provided evidence that judgments were being made based on appearance-activated biases, not just on the actual speech heard.

Ultimately, appearance played little role in the judgments of L2 speech and, based on social attitudes questionnaire responses, raters exhibited generally positive attitudes toward immigrants overall. What did surface, however, consistent with group bias theories, is a clear hierarchy of language groups, such that certain L2 speaker groups were consistently downgraded in linguistic measures, status and solidarity traits, and citizenship status in comparison to native English speakers. These findings suggest that perhaps a more nuanced approach to group bias is needed

in investigation of native-born residents' attitudes toward immigrants. While Canadians might not feel universally threatened by all immigrants, certain L2 speaker groups appear to be judged differently in terms of social dominance and belonging. While it is not clear from this research precisely what promotes this hierarchy, a closer look at the demographics and integration of certain L2 groups (Lindsay, 2007) suggests support for social identity theory (Tajfel, 1972), in that groups that have more in common with the raters (e.g., language spoken at home, religion practiced, proficiency in English and/or French) are rated more favourably across dimensions. Therefore, one way to gain deeper understanding of intergroup dynamics in Canada is to shift focus and look at the relationship between native-born residents and L2 speakers from a different angle. In other words, instead of focusing on rater factors in group biases, we might take a closer look at the L2 speakers themselves through theories that explore the role of language proficiency in acceptance by a host community, for instance, or through various assimilation theories that explore proximal (individual, group) or distal (institutional, government, economic) assimilation (Alba & Nee, 2003), as well as segmented assimilation (Portes & Zhou, 1993), which posits that different groups have different assimilation processes that are largely influenced by their social environment. A working hypothesis underlying this future work might be that issues of language proficiency play a key role in determining the social environment for a given L2 speaker group. Through investigation of the various ways in which these groups incorporate themselves into Canadian society, we might better understand why native-born residents' attitudes toward L2 speakers converge or diverge.

Societal Implications

Another reason why group bias might not have been as prominent as expected is because attitudes and actions are two different things. For instance, raters may behave differently in a lab

environment or when completing online tasks in their own homes (as was the case in this study) than they would in everyday interactions. Here, raters were aware that their answers would be reviewed and had plenty of time to consider their responses. As a result, the current studies potentially present a softened view of how native-born residents truly judge L2 speakers. In addition, even when negative attitudes are present, there is no guarantee that those biases will lead to discriminatory behaviour (Krieger, 1998), and *action* toward L2 speakers was not studied as part of this project. Furthermore, raters might not hold complete awareness of their biases or their awareness might be context-dependent. On the social attitudes questionnaire, for instance, raters might have wanted to express general acceptance of immigrants as part of a conscious, controlled action (Wilson et al., 2000), but raters' everyday interactions with and their treatment of immigrants are perhaps not so guarded. Even through matched-guise studies, which explore biases that are arguably less explicit than those that would emerge through direct questioning about attitudes toward immigrants, participants are often aware of the study's objective (Zheng & Samuel, 2017). Therefore, it would be important to explore native-born residents' attitudes toward specific L2 groups through more implicit measures (e.g., Green et al., 2007) or through measures of direct observation, such as through monitoring of discriminatory complaints in human rights cases (Munro, 2003).

In addition to exploring various ways to document biases and discrimination, it is important to consider what steps should be taken once negative behaviour is identified. This dissertation project was originally designed to include a mitigation component in the form of a storytelling event involving both native-born residents and immigrants that would be carried out in the context (Montreal or Calgary) where negative biases were most prominent. The COVID-19 pandemic temporarily removed the possibility of such prolonged face-to-face interactions, but

their power in forming stronger intergroup relationships should still be considered. Although general negative biases toward immigrants did not surface in either context, the hierarchy of L2 speaker group signals an area of potential improvement for intergroup relations in Canada. Here, we might first look to social network theory. Overall, the raters in this study had fairly diverse social networks representing over 40 different language backgrounds and 11 religions. This exposure most likely explains raters' generally positive attitudes toward immigrants and supports social network theory (Eisenstadt, 1961), which suggests that social contact can build familiarity, lessen social distance, and foster intergroup acceptance. Indeed, learning about someone else's culture through regular interaction can lessen any potential threat that a person might pose to a majority group's identity and can help mitigate the group's biases toward that person (Escandell & Ceobanu, 2009; Hewstone, 2015; Tropp & Pettigrew, 2005). However, there might be a limit to this benefit, such that with a fair amount of diversity in any given network, its positive effect on mitigating general negative attitudes reaches a plateau. That is, greater network diversity might not always lead to greater acceptance. Therefore, promoting expansion of a network as mitigation is mostly useful among those who have limited networks and for targeted exposure to specific L2 speaker groups. Along these lines, considering the variable judgments that emerged about the speaker groups in this dissertation, it would still be important to encourage the expansion of native-born Canadian residents' social networks through targeted intergroup contact (Allport, 1954), particularly among groups that rarely interact, in an attempt to lessen the hurdles of belonging already experienced by L2 speakers. In fact, capitalizing on Canada's support for multilingualism by encouraging residents to study the languages of heavily marginalized L2 speaker groups could also play a role in building intergroup relationships (Genesee & Gándara, 1999).

Limitations and Suggestions for Future Research

While the overall scope of this dissertation has enabled a comprehensive view of social attitudes toward L2 speakers, there are many ways in which future investigations could be improved or expanded. First, the two contexts chosen are cities rich in diversity, so there was already a form of mitigation in place that was likely to promote more positive social attitudes toward L2 speakers. Future research should extend to isolated and rural areas as well as to metropolitan environments that do not host large immigrant populations. In doing so, the role of social network exposure would also warrant reconsideration. In addition, the speech samples used across studies were of one male and one female per speaker group. Larger groups of male, female, and non-binary individuals would allow for a more representative population of L2 speakers and, as a result, better generalizability of the findings. A larger group of speakers per L2 group (and perhaps fewer groups) would also enrich our understanding, particularly of attitudes toward speakers often overlooked in matched-guise research (Lee & Bailey, 2022). This project also employed speech samples that were gathered as part of a mock interview and that were likely more formal than most residents' interactions with L2 speakers, so it would be important to conduct similar investigations using conversational speech. Furthermore, it is impossible to gauge the role that individual characteristics (e.g., personality, confidence, approach) and speech characteristics (e.g., rate, pausing, pitch) played in judgments of the speakers in this investigation, and language proficiency—although self-reported by all speakers as intermediate to high intermediate—was not specifically assessed. Including these elements in

pre-rating tasks and considering their role in rating outcomes could allow for analyses that focus on characteristics indicative of a particular language group rather than of an individual speaker.

There were also a few task-related limitations worth noting. For example, the tasks were presented in such a way that background information (including social network data) was collected first, followed by speech ratings, then by the social attitudes questionnaire. Given the fact that questioning raters about their attitudes toward immigrants could potentially *engage* biases, it would be important to experiment with different task configurations, perhaps even by using two different configurations among a large group of raters to explore the effects of task order. Furthermore, there was a potential mismatch between the presentation of the social attitudes questionnaire and the speech rating task. While the questionnaire explored attitudes toward immigrants, the speakers were identified simply as L2 speakers (although all of the L2 speakers were, in fact, born outside Canada). Future research should ensure alignment of questionnaire terminology and speaker presentation and perhaps use this pairing to explore social attitudes toward a wider range of L2 speaker out-groups (e.g., newly-arrived refugees, international students, migrant farm workers, visiting professors/researchers). There were also no time constraints attached to any task in this investigation, which means that participants had unlimited time to consider their responses and perhaps put forth a politically correct judgment. Therefore, it would be important to consider the use of implicit association tasks and to explore judgments completed within a limited timeframe. Finally, a surprising finding in terms of how native-born residents view the citizenship of various L2 speaker groups should serve as a call to action, as this project appears to be the first to explore this particular aspect of social attitudes. Future research should involve not only native-born residents' perceptions of L2 speaker citizenship, but also those speakers' perceptions of their own belonging. This could include

various task designs in which citizenship categories are defined or perhaps limited to a simple choice between whether the speaker is a Canadian citizen or not.

Conclusion

The overall goal of this dissertation was to investigate native-born residents' speech judgments and social attitudes toward L2 speakers of English in two representative Canadian contexts. Ultimately, raters in Montreal and Calgary demonstrated similar attitudes and were not swayed by visual suggestions of speaker ethnicity or religion in their ratings of linguistic or extralinguistic attributes of L2 speakers. However, results showed that not all L2 speaker groups are seen in the same light, revealing a rater preference for speakers who are less accented and easier to understand that led to more favourable judgments of those speakers across rated dimensions. Encouragingly, raters shared generally positive attitudes toward immigrants overall and, despite subtle (but significant) differences in their social networks, neither their interaction with L2 speakers nor their social attitudes were linked to their ratings. These findings support our understanding of the saliency of L2 speech and of the roles of context and rater background in the formation of social attitudes.

References

- Adserà, A., & Pytliková, M. (2016). Language and migration. In *The Palgrave Handbook of Economics and Language* (pp. 342–372). Palgrave Macmillan, London. https://doi.org/10.1007/978-1-137-32505-1_13
- Agrawal, S., & Kurtz, N. (2019). Ethnic spatial segmentation in immigrant destinations—Edmonton and Calgary. *Journal of International Migration and Integration*, 20, 199–222. <https://doi.org/10.1007/s12134-018-0604-y>
- Ahmed, S., Chen, V. H. H., & Chib, A. I. (2021). Xenophobia in the time of a pandemic: social media use, stereotypes, and prejudice against immigrants during the COVID-19 crisis. *International Journal of Public Opinion Research*, 33, 637–653. <https://doi.org/10.1093/ijpor/edab014>
- Akehurst, L., Arnhold, A., Figueiredo, I., Turtle, S., & Leach, A. M. (2018). Investigating deception in second language speakers: Interviewee and assessor perspectives. *Legal and Criminological Psychology*, 23(2), 1–22. <https://doi.org/10.1111/lcrp.12127>
- Al Ramiah, A., & Hewstone, M. (2013). Intergroup contact as a tool for reducing, resolving, and preventing intergroup conflict: evidence, limitations, and potential. *American Psychologist*, 68, 527–542. <https://doi.org/10.1037/a0032603>
- Alba, R., & Nee, V. (2003). Assimilation theory, old and new. In R. Alba & V. Nee (Eds.), *Remaking the American mainstream: Assimilation and contemporary immigration* (pp. 17–66). Cambridge, MA: Harvard University Press. <https://doi.org/10.4159/9780674020115-002>
- Albarello, F., & Rubini, M. (2018). Linguistic discrimination toward Roma: Can intergroup threat enhance bias?. *Journal of Language and Social Psychology*, 37, 350–364. <https://doi.org/10.1177/0261927x17725880>

Allport, G. W. (1954). *The nature of prejudice*. Boston: Addison-Wesley.

<https://doi.org/10.2307/1418507>

Angus Reid (2017). *Religious trends: Led by Quebec, number of Canadians holding favourable views of various religions increases*. <https://angusreid.org>

Atanackovic, J., & Bourgeault, I. L. (2013). The employment and recruitment of immigrant care workers in Canada. *Canadian Public Policy*, 39, 335–350. <https://doi.org/10.3138/cpp.39.2.335>

Atwell Seate, A., & Mastro, D. (2016). Media's influence on immigration attitudes: An intergroup threat theory approach. *Communication Monographs*, 83, 194–213.

<https://doi.org/10.1080/03637751.2015.1068433>

Aujla-Bhullar, S. K. (2020). Crowns and cages: A Sikh woman's reflections of the Sikh community in Canada. *Cultural and Pedagogical Inquiry*, 12(1), 73–82. <https://doi.org/10.18733/cpi29532>

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

Baese-Berk, M. M., McLaughlin, D. J., & McGowan, K. B. (2020). Perception of non-native speech. *Language and Linguistics Compass*, 14(7), e12375. <https://doi.org/10.1111/lnc3.12375>

Baquiran, C. L. C., & Nicoladis, E. (2020). A doctor's foreign accent affects perceptions of competence. *Health Communication*, 35, 726–730.

<https://doi.org/10.1080/10410236.2019.1584779>

Barona, D. B. (2008). Native and non-native speakers' perceptions of non-native accents. *Language and Literature Journal*, 3(2). <http://ojs.gc.cuny.edu/index.php/lljournal/article/viewArticle/430/428>.

Bartko, K., Offin, S., & Hughes, J. (2021, Oct. 26). *Racist graffiti in Calgary's southwest considered a hate crime: police*. Global News. <https://globalnews.ca/news/8325615/sikh-society-of-calgary-gurdwara-graffiti-october-2021/>

- Bastian, B., & Haslam, N. (2008). Immigration from the perspective of hosts and immigrants: Roles of psychological essentialism and social identity. *Asian Journal of Social Psychology*, *11*, 127–140. <https://doi.org/10.1111/j.1467-839x.2008.00250.x>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*. <https://doi.org/10.18637/jss.v067.i01>
- Beaulieu, S., Bejarano, J., French, L. M., & Reinke, K. (2022). Professional identities of French Lx economic immigrants: Perceptions from a local French-speaking community. *Languages*, *7*(2), 1–15. <https://doi.org/10.3390/languages7020140>
- Benish-Weisman, M., & Horenczyk, G. (2010). Cultural identity and perceived success among Israeli immigrants: An emic approach. *International Journal of Intercultural Relations*, *34*, 516–526. <https://doi.org/10.1016/j.ijintrel.2009.11.010>
- Berry, J. W. (2013). Research on multiculturalism in Canada. *International Journal of Intercultural Relations*, *37*, 663–675. <https://doi.org/10.1016/j.ijintrel.2013.09.005>
- Berry, J. W. (2006). Mutual attitudes among immigrants and ethnocultural groups in Canada. *International journal of intercultural relations*, *30*, 719–734. <https://doi.org/10.1016/j.ijintrel.2006.06.004>
- Berry, J. W., & Hou, F. (2017). Acculturation, discrimination and wellbeing among second generation of immigrants in Canada. *International Journal of Intercultural Relations*, *61*, 29–39. <https://doi.org/10.1016/j.ijintrel.2017.08.003>
- Blake, R. (2019). Citizenship, national identity, and the search for stability in Canada. *Acta Universitatis Carolinae Studia Territorialia*, *19*(2), 11–38. <https://doi.org/10.14712/23363231.2020.2>
- Blalock, H.M (1967) *Towards a theory of minority group relations*. Wiley. <https://doi.org/10.2307/3002081>

- Blalock, H. M. (1956). Economic discrimination and Negro increase. *American Sociological Review*, *21*, 584–588. <https://doi.org/10.2307/2089093>
- Bloemraad, I. (2000). Citizenship and immigration a current review. *Journal of International Migration and Integration*, *1*(1), 9–37. <https://doi.org/10.1007/s12134-000-1006-4>
- Blumer, H. (1958). Race prejudice as a sense of group position. *Pacific Sociological Review*, *1*, 3–7. https://doi.org/10.1007/978-1-349-26403-2_3
- Bobowik, M., Benet-Martínez, V., & Repke, L. (2021). Ethnocultural diversity of immigrants' personal social networks, bicultural identity integration and global identification. *International Journal of Psychology*, *57*. <https://doi.org/10.1002/ijop.12814>
- Bochner, S. (1982). The social psychology of cross-cultural relations. *Cultures in Contact: Studies in Cross-Cultural Interaction*, *1*, 5–44. <https://doi.org/10.1016/b978-0-08-025805-8.50008-1>
- Boschken, H. L. (2003). Global cities, systemic power, and upper-middle-class influence. *Urban Affairs Review*, *38*, 808–830. <https://doi.org/10.1177/1078087403038006003>
- Bourdieu, P. (1977). The economics of linguistic exchanges. *Social Science Information*, *16*, 645–668. <https://doi.org/10.1177/053901847701600601>
- Bourhis, R. Y. (2019). Evaluating the impact of Bill 101 on the English-speaking communities of Quebec. *Language Problems and Language Planning*, *43*, 198–229. <https://doi.org/10.1075/lplp.00042.bou>
- Bourhis, R. Y., & Giles, H. (1977). The language of intergroup distinctiveness. *Language, Ethnicity, and Intergroup Relations*, *13*, 119–135.
- Bourhis, R. Y., Montaruli, E., El-Geledi, S., Harvey, S. P., & Barrette, G. (2010). Acculturation in multiple host community settings. *Journal of Social Issues*, *66*, 780–802. <https://doi.org/10.1111/j.1540-4560.2010.01675.x>

- Bousmah, I., Grenier, G., & Gray, D. M. (2021). Linguistic distance, languages of work and wages of immigrants in Montreal. *Journal of Labor Research*, 42(1), 1–28.
<https://doi.org/10.1007/s12122-020-09316-1>
- Boyd, M. (2009, October 22). *Official language proficiency and the civic participation of immigrants*. [Paper presentation]. Metropolis Language Matters Symposium, Ottawa, Canada.
- Boyd, M., & Alboim, N. (2012). Managing international migration: The Canadian case. In D. Rodriguez-Garcia (Ed.), *Managing immigration and diversity in Canada: A transatlantic dialogue in the new age of migration* (pp. 123–50). McGill-Queens University Press.
- Boyd, M., & Schellenberg, G. (2007). Re-accreditation and the occupations of immigrant doctors and engineers. *Canadian Social Trends*, 84, 2–10.
- Brennan, K. M., & London, A. S. (2001). Are religious people nice people? Religiosity, race, interview dynamics, and perceived cooperativeness. *Sociological Inquiry*, 71(2), 129–144.
<https://doi.org/10.1111/j.1475-682x.2001.tb01105.x>
- Brosseau, L., & Dewing, M. (2018). *Canadian multiculturalism*. Library of Parliament.
<https://lop.parl.ca/staticfiles/PublicWebsite/Home/ResearchPublications/BackgroundPapers/PDF/2009-20-e.pdf>
- Brownstein, B. (2021, April 16). *Anti-Asian racism has immigrant rethinking life here*. Montreal Gazette. <https://montrealgazette.com/opinion/brownstein-anti-asian-racism-has-immigrant-rethinking-life-here>
- Canadian Civil Liberties Association. (2022). *Why Bill-21 is an issue*. <https://ccla.org/major-cases-and-reports/bill-21/>

- Cantone, J. A., Martinez, L. N., Willis-Esqueda, C., & Miller, T. (2019). Sounding guilty: How accent bias affects juror judgments of culpability. *Journal of Ethnicity in Criminal Justice, 17*, 228–253. <https://doi.org/10.1080/15377938.2019.1623963>
- Cantone, J. A., & Wiener, R. L. (2017). Religion at work: Evaluating hostile work environment religious discrimination claims. *Psychology, Public Policy, and Law, 23*, 351–366. <https://doi.org/10.1037/law0000132>
- Carlson, H. K., & McHenry, M. A. (2006). Effect of accent and dialect on employability. *Journal of Employment Counseling, 43*, 70–83. <https://doi.org/10.1002/j.2161-1920.2006.tb00008.x>
- Chand, M., & Tung, R. L. (2019). Skilled immigration to fill talent gaps: A comparison of the immigration policies of the United States, Canada, and Australia. *Journal of International Business Policy, 2*, 333–355. <https://doi.org/10.1057/s42214-019-00039-4>
- Charland, W. (1977). Multilingualism and the CBC Mandate: An example of ineffectual regulation. *Dalhousie Law Journal, 4*, 166–188.
- Cheng, C. (1997). Are Asian American employees a model minority or just a minority?. *The Journal of Applied Behavioral Science, 33*, 277–290. <https://doi.org/10.1177/00218863973333002>
- Chuah, S. H., Gächter, S., Hoffmann, R., & Tan, J. H. (2016). Religion, discrimination and trust across three cultures. *European Economic Review, 90*, 280–301. <https://doi.org/10.1016/j.euroecorev.2016.03.008>
- Clark, T. (2006). Language as social capital. *Applied Semiotics, 8*(18), 29–41.
- Creese, G. (2010). Erasing English language competency: African migrants in Vancouver, Canada. *Journal of International Migration and Integration, 11*, 295–313. <https://doi.org/10.1007/s12134-010-0139-3>

- CROP (2017). Les Canadiens, le populisme et la xénophobie. *Radio-Canada: Rapport présenté par CROP*.
- Croucher, S., Galy-Badenas, F., & Routsalainen, M. (2014). Host culture acceptance, religiosity, and the threat of Muslim immigration: An integrated threat analysis in Spain. *Journal of Intercultural Communication, 15*(35), 1–16.
- Danso, R. K., & Grant, M. R. (2000). Access to housing as an adaptive strategy for immigrant groups: Africans in Calgary. *Canadian Ethnic Studies Journal, 32*(3), 19–25.
- Da Silva, E., McLaughlin, M., & Richards, M. (2007). Bilingualism and the globalized new economy: The commodification of language and identity. In M. Heller (Ed.), *Bilingualism: A social approach* (pp. 183–206). Palgrave Macmillan. https://doi.org/10.1057/9780230596047_9
- Dechief, D., & Oreopoulos, P. (2012). *Why do some employers prefer to interview Matthew but not Samir? New evidence from Toronto, Montreal and Vancouver* (Working Paper No. 95). <https://doi.org/10.2139/ssrn.2018047>
- Deng, J., & Walker, G. J. (2007). Chinese acculturation measurement. *Canadian Ethnic Studies, 39*(1-2), 187–217. doi:10.1353/ces.0.0008.
- Department of Canadian Heritage (2012). *Official languages annual report, 2010–11*. Her Majesty the Queen in Right of Canada. https://publications.gc.ca/collections/collection_2012/pc-ch/CH10-2011-1-eng.pdf
- Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. John Benjamins.
- Derwing, T. M., & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching, 42*, 476–490. <https://doi.org/10.1075/llt.42>

- Derwing, T. M., Munro, M. J., Mulder, M., & Abbott, M. (2010). *An examination of the Canadian language benchmark data from the citizenship language survey*. Citizenship and Immigration Canada. <http://www.cic.gc.ca/english/resources/research/language-benchmark/index.asp>
- Derwing, T. M., & Waugh, E. (2012). *Language skills and the social integration of Canada's adult immigrants* (IRPP Study 31). Institute for Research on Public Policy. <https://irpp.org/research-studies/language-skills-and-the-social-integration-of-canadas-adult-immigrants/>
- de Souza, L. E. C., Pereira, C. R., Camino, L., de Lima, T. J. S., & Torres, A. R. R. (2016). The legitimizing role of accent on discrimination against immigrants. *European Journal of Social Psychology, 46*, 609–620. <https://doi.org/10.1002/ejsp.2216>
- Devos, T., & Banaji, M. R. (2005). American = White? *Journal of Personality and Social Psychology, 88*, 447–466. <https://doi.org/10.1037/0022-3514.88.3.447>
- Devos, T., Gavin, K., & Quintana, F. J. (2010). Say “adios” to the American dream? The interplay between ethnic and national identity among Latino and Caucasian Americans. *Cultural Diversity & Ethnic Minority Psychology, 16*(1), 37–49. <https://doi.org/10.1037/a0015868>
- Dhont, K., Van Hiel, A., & Hewstone, M. (2014). Changing the ideological roots of prejudice: Longitudinal effects of ethnic intergroup contact on social dominance orientation. *Group Processes & Intergroup Relations, 17*(1), 27–44. <https://doi.org/10.1177/1368430213497064>
- Dion, K. L. (2001). Immigrants' perceptions of housing discrimination in Toronto: The Housing New Canadians Project. *Journal of Social Issues, 57*, 523–539. <https://doi.org/10.1111/0022-4537.00227>
- Doucerein, M., Varnaamkhaastil, R., Segalowitz, N., & Ryder, A. (2015). Second language social networks and communication-related acculturative stress: The role of interconnectedness. *Frontiers in Psychology, 6*, 1–12. <https://doi.org/10.3389/fpsyg.2015.01111>

- Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology*, 82(1), 62–68.
<https://doi.org/10.1037/0022-3514.82.1.62>
- Dragojevic, M., Berglund, C., & Blauvelt, T. K. (2018). Figuring out who's who: The role of social categorization in the language attitudes process. *Journal of Language and Social Psychology*, 37(1), 28–50. <https://doi.org/10.1177/0261927x17706942>
- Dragojevic, M., & 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*, 42, 396–420. <https://doi.org/10.1111/hcre.12079>
- Dragojevic, M., & Goatley-Soan, S. (2022). Americans' attitudes toward foreign accents: Evaluative hierarchies and underlying processes. *Journal of multilingual and multicultural development*, 43(2), 167–181. <https://doi.org/10.1080/01434632.2020.1735402>
- Eberhardt, J. L. (2020). *Biased: Uncovering the hidden prejudice that shapes what we see, think, and do*. Penguin Books.
- Edmonston, B. (2021). Asian immigrant advancement in Canada. *Indian Journal of Population and Development*, 1(2), 145–184.
- Eisenstadt, S. N. (1961). Anthropological studies of complex societies. *Current Anthropology*, 2, 201–222. <https://doi.org/10.1086/200188>
- Ellison, C. G. (1992). Are religious people nice people? Evidence from the National Survey of Black Americans. *Social Forces*, 71, 411–430. <https://doi.org/10.2307/2580017>
- Escandell, X., & Ceobanu, A. M. (2009). When contact with immigrants matters: threat, interethnic attitudes and foreigner exclusionism in Spain's Comunidades Autónomas. *Ethnic and Racial Studies*, 32(1), 44–69. <https://doi.org/10.1080/01419870701846924>

- Esses, V. M. (2021). Prejudice and discrimination toward immigrants. *Annual Review of Psychology*, 72, 503–531. <https://doi.org/10.1146/annurev-psych-080520-102803>
- Esses, V. M., & Abelson, D. E. (Eds.). (2017). *Twenty-first-century immigration to North America: Newcomers in turbulent times*. McGill-Queen's Press-MQUP. <https://doi.org/10.2307/j.ctt1w6tdzn>
- Esses, V. M., Jackson, L. M., & Armstrong, T. L. (1998). Intergroup competition and attitudes toward immigrants and immigration: An instrumental model of group conflict. *Journal of Social Issues*, 54, 699–724. <https://doi.org/10.1111/j.1540-4560.1998.tb01244.x>
- Esses, V. M., Wagner, U., Wolf, C., Preiser, M., & Wilbur, C. J. (2006). Perceptions of national identity and attitudes toward immigrants and immigration in Canada and Germany. *International Journal of Intercultural Relations*, 30, 653–669. <https://doi.org/10.1016/j.ijintrel.2006.07.002>
- European Social Survey. (2015). *Round 7 module on attitudes towards immigration and their antecedents: Questions from the round 7 pilot questionnaire*. Centre for Comparative Social Surveys, City University London.
- Fayer, J. M., & Krasinski, E. (1987). Native and nonnative judgments of intelligibility and irritation. *Language Learning*, 37, 313–326. <https://doi.org/10.1111/j.1467-1770.1987.tb00573.x>
- Fenn, K. (2021, June 9). *These Muslim women say harassment because of their faith is constant and relentless*. CBC Radio. <https://www.cbc.ca/radio/thecurrent/the-current-for-june-9-2021-1.6058842/these-muslim-women-say-harassment-because-of-their-faith-is-constant-and-relentless-1.6059312>
- Ford, C. E. (1984). The influence of speech variety on teachers' evaluation of students with comparable academic ability. *TESOL Quarterly*, 18(1), 25–40. <https://doi.org/10.2307/3586333>

- Foucart, A., Costa, A., Morís-Fernández, L., & Hartsuiker, R. J. (2020). Foreignness or processing fluency? On understanding the negative bias toward foreign-accented speakers. *Language Learning, 70*, 974–1016. <https://doi.org/10.1111/lang.12413>
- Fournier, P., Cutler, F., Soroka, S., & Stolle, D. (2015). *Canadian Election Study (CES)*. <http://ces-ec.arts.ubc.ca/>
- Frumkin, L. A., & Stone, A. (2020). Not all eyewitnesses are equal: Accent status, race and age interact to influence evaluations of testimony. *Journal of Ethnicity in Criminal Justice, 18*(2), 123–145. <https://doi.org/10.1080/15377938.2020.1727806>
- Fuertes, J. N., Gottdiener, W. H., Martin, H., Gilbert, T. C., & Giles, H. (2012). A meta-analysis of the effects of speakers' accents on interpersonal evaluations. *European Journal of Social Psychology, 42*(1), 120–133. <https://doi.org/10.1002/ejsp.862>
- Gaither, S. E., Cohen-Goldberg, A. M., Gidney, C. L., & Maddox, K. B. (2015). Sounding black or white: Priming identity and biracial speech. *Frontiers in Psychology, 6*, 1–11. <https://doi.org/10.3389/fpsyg.2015.00457>
- Galupo, M. P., & Gonzalez, K. A. (2013). Friendship values and cross-category friendships: Understanding adult friendship patterns across gender, sexual orientation and race. *Sex Roles, 68*, 779–790. <https://doi.org/10.1007/s11199-012-0211-x>
- Genesee, F., & Gándara, P. (1999). Bilingual education programs: A cross-national perspective. *Journal of Social Issues, 55*, 665–685. <https://doi.org/10.1111/0022-4537.00141>
- George, G., & Selimos, E. D. (2019). Searching for belonging and confronting exclusion: A person-centred approach to immigrant settlement experiences in Canada. *Social Identities, 25*(2), 125–140. <https://doi.org/10.1080/13504630.2017.1381834>

- Ghanem, R., & Kang, O. (2021). ESL students' reverse linguistic stereotyping of English teachers. *ELT Journal*, 75, 330–340. <https://doi.org/10.1093/elt/ccab011>
- Gilligan, M. (2021, June 16). *Mother, kids verbally accosted in downtown Calgary in hate-motivated incident*. Global News. <https://globalnews.ca/news/7954247/calgary-muslim-woman-berated-hate-motivated/>
- Godley, J. (2018). Everyday discrimination in Canada: Prevalence and patterns. *Canadian Journal of Sociology*, 43(2), 111–142. <https://doi.org/10.29173/cjs29346>
- Goldin, C., & Rouse, C. (2000). Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American Economic Review*, 90, 715–741. <https://doi.org/10.3386/w5903>
- Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *American Psychologist*, 59(2), 93–104. <https://doi.org/10.1037/0003-066x.59.2.93>
- Government of Canada (2022). *Immigration, refugees and citizenship Canada*. <https://www.canada.ca/en/immigration-refugees-citizenship.html>
- Government of Canada (2018). *Inclusion and respect for diversity*. https://www.international.gc.ca/world-monde/issues_developpement-enjeux_developpement/human_rights-droits_homme/inclusion_respect.aspx?lang=eng
- Government of Canada (2022). *What are the requirements for becoming a Canadian citizen?* <https://www.cic.gc.ca/english/helpcentre/answer.asp?qnum=355&top=5>
- Graham, J. H. and Thurston, W. E. (2005). Overcoming adversity: Resilience and coping mechanisms developed by recent immigrant women living in the inner-city of Calgary, Alberta. *Women's Health and Urban Life*, 4, 63–80.

- Gravelle, T. B. (2018). Partisanship, local context, group threat, and Canadian attitudes towards immigration and refugee policy. *Migration Studies*, 6, 448–467.
<https://doi.org/10.1093/migration/mnx058>
- Green, A. R., Carney, D. R., Pallin, D. J., Ngo, L. H., Raymond, K. L., Iezzoni, L. I., & Banaji, M. R. (2007). Implicit bias among physicians and its prediction of thrombolysis decisions for black and white patients. *Journal of General Internal Medicine*, 22, 1231–1238.
<https://doi.org/10.1007/s11606-007-0258-5>
- Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, 25, 515–530.
<https://doi.org/10.1080/20445911.2013.796377>
- Grenke, A. (2018). *German Canadians: Community formation, transformation and contribution to Canadian life*. Trafford Publishing.
- Guboglo, M. N. (1974). Socioethnic consequences of bilingualism. *Soviet Sociology*, 13(1–2), 93–113.
<https://doi.org/10.2753/sor1061-015413010293>
- Guo, S., & Guo, Y. (2016). Immigration, integration and welcoming communities: neighbourhood-based initiative to facilitate the integration of newcomers in Calgary. *Canadian Ethnic Studies*, 48(3), 45–67. <https://doi.org/10.1353/ces.2016.0025>
- Hall, D. L., Matz, D. C., & Wood, W. (2010). Why don't we practice what we preach? A meta-analytic review of religious racism. *Personality and Social Psychology Review*, 14(1), 126–139.
<https://doi.org/10.1177/1088868309352179>
- Hansen, K., & Dovidio, J. F. (2016). Social dominance orientation, nonnative accents, and hiring recommendations. *Cultural Diversity and Ethnic Minority Psychology*, 22(4), 1–8.
<https://doi.org/10.1037/cdp0000101>

- Hansen, K., Rakić, T., & Steffens, M. C. (2017a). Competent and warm? How mismatching appearance and accent influence first impressions. *Experimental Psychology*, *64*(1), 27–36.
<https://doi.org/10.1027/1618-3169/a000348>
- Hansen, K., Steffens, M. C., Rakić, T., & Wiese, H. (2017b). When appearance does not match accent: neural correlates of ethnicity-related expectancy violations. *Social Cognitive and Affective Neuroscience*, *12*, 507–515. <https://doi.org/10.1093/scan/nsw148>
- Harding, A., & Blokland, T. (2014). *Urban theory: a critical introduction to power, cities and urbanism in the 21st century*. Sage.
- Harell, A. (2010). Political tolerance, racist speech, and the influence of social networks. *Social Science Quarterly*, *91*, 724–740. <https://doi.org/10.1111/j.1540-6237.2010.00716.x>
- Harell, A., Soroka, S., Iyengar, S., & Valentino, N. (2012). The impact of economic and cultural cues on support for immigration in Canada and the United States. *Canadian Journal of Political Science*, *45*, 499–530. <https://doi.org/10.1017/s0008423912000698>
- Harrison, G. (2013). “Oh, you've got such a strong accent”: Language identity intersecting with professional identity in the human services in Australia. *International Migration*, *51*(5), 192–204. <https://doi.org/10.1111/imig.12005>
- Herrnstein, R. J., & Murray, C. (1994). *The Bell Curve: Intelligence and class structure in American life*. The Free Press.
- Hewstone, M. (2015). Consequences of diversity for social cohesion and prejudice: The missing dimension of intergroup contact. *Journal of Social Issues*, *71*, 417–438.
<https://doi.org/10.1111/josi.12120>

- Hiebert, D. (2016). What's so special about Canada? Understanding the resilience of immigration and multiculturalism. *Migration Policy Institute*. <https://www.migrationpolicy.org/research/whats-so-special-about-canada-understanding-resilience-immigration-and-multiculturalism>
- Hodson, G., Hewstone, M., & Swart, H. (2013). Epilogue and future directions. In G. Hodson & M. Hewstone (Eds.), *Advances in intergroup contact* (pp. 262–297). Psychology Press.
- Hogg, M. A. (2006). Social identity theory. In P. J. Burke (Ed.), *Contemporary social psychological theories* (pp. 111–136). Stanford University Press. <https://doi.org/10.1515/9780804768047-008>
- Holladay, S. J., & Coombs, W. T. (1994). Speaking of visions and visions being spoken: An exploration of the effects of content and delivery on perceptions of leader charisma. *Management Communication Quarterly*, 8(2), 165–189. <https://doi.org/10.1177/0893318994008002002>
- Hosoda, M., & Stone-Romero, E. (2010). The effects of foreign accents on employment-related decisions. *Journal of Managerial Psychology*, 25(2), 113–132. <https://doi.org/10.1108/02683941011019339>
- Hou, F., & Picot, W. G. (2019). *Trends in the citizenship rate among new immigrants to Canada*. Statistics Canada. <https://doi.org/10.2139/ssrn.2012582>
- Hu, G., & Lindemann, S. (2009). Stereotypes of Cantonese English, apparent native/non-native status, and their effect on non-native English speakers' perception. *Journal of Multilingual and Multicultural Development*, 30, 253–269. <https://doi.org/10.1080/01434630802651677>
- Immigration, Francisation et Intégration Québec. (2019). *Recherche et statistiques*. <http://www.mifi.gouv.qc.ca/fr/recherches-statistiques/index.html>
- Ipsos MORI (2016). *Global views on immigration and the refugee crisis*. <https://www.ipsos.com/en/global-views-immigration-and-refugee-crisis>

- Jian, G. (2012). Does culture matter? An examination of the association of immigrants' acculturation with workplace relationship quality. *Management Communication Quarterly*, 26, 295–321.
<https://doi.org/10.1177/0893318912440178>
- Johnston, H. (2005). Sikhs in Canada. In M. Ember, C. R. Ember, & I. Skoggard (Eds.), *Encyclopedia of diasporas: Immigrant and refugee cultures around the world*. (pp. 1075–1083). Springer Science and Business Media, Inc. https://doi.org/10.1007/978-0-387-29904-4_110
- Joshee, R., & Thomas, M. (2017). Multicultural and citizenship education in Canada: Slow peace as an alternative to social cohesion. In J. A. Banks (Ed.), *Citizenship education and global migration: Implications for theory, research, and teaching* (pp. 91–106). American Educational Research Association. <https://doi.org/10.2307/j.ctv138wrq5.12>
- Kang, O., & Rubin, D. L. (2009). Reverse linguistic stereotyping: Measuring the effect of listener expectations on speech evaluation. *Journal of Language and Social Psychology*, 28, 441–456.
<https://doi.org/10.1177/0261927x09341950>
- Kang, O., Rubin, D., & Lindemann, S. (2015). Mitigating US undergraduates' attitudes toward international teaching assistants. *Tesol Quarterly*, 49, 681-706. <https://doi.org/10.1002/tesq.192>
- Kang, O., Rubin, D., & Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *The Modern Language Journal*, 94, 554–566.
<https://doi.org/10.1111/j.1540-4781.2010.01091.x>
- Kang, O., & Yaw, K. (2021). Social judgement of L2 accented speech stereotyping and its influential factors. *Journal of Multilingual and Multicultural Development*, 1–16.
<https://doi.org/10.1080/01434632.2021.1931247>

- Kemmelmeier, M., & Winter, D. G. (2000). Putting threat into perspective: Experimental studies on perceptual distortion in international conflict. *Personality and Social Psychology Bulletin*, 26, 795–809. <https://doi.org/10.1177/0146167200269005>
- Kent, J. (2022). Can urban fabric encourage tolerance? Evidence that the structure of cities influences attitudes toward migrants in Europe. *Cities*, 121. <https://doi.org/10.1016/j.cities.2021.103494>
- Kim, R. Y. (2011). Religion and ethnicity: Theoretical connections. *Religions*, 2, 312–329. <https://doi.org/10.3390/rel2030312>
- Kinzler, K. D., Shutts, K., DeJesus, J., & Spelke, E. S. (2009). Accent trumps race in guiding children's social preferences. *Social Cognition*, 27, 623–634. <https://doi.org/10.1521/soco.2009.27.4.623>
- Kircher, R. (2012). How pluricentric is the French language? An investigation of attitudes towards Quebec French compared to European French. *Journal of French Language Studies*, 22, 345–370. <https://doi.org/10.1017/s0959269512000014>
- Klitmøller, A., Schneider, S. C., & Jonsen, K. (2015). Speaking of global virtual teams: Language differences, social categorization and media choice. *Personnel Review*, 44, 270–285. <https://doi.org/10.1108/pr-11-2013-0205>
- Krieger, L. H. (1998). Civil rights perestroika: Intergroup relations after affirmative action. *California Law Review*, 86, 1251–1329. <https://doi.org/10.2307/3481107>
- Kutlu, E. (2020). Now you see me, now you mishear me: Raciolinguistic accounts of speech perception in different English varieties. *Journal of Multilingual and Multicultural Development*, 1–15. <https://doi.org/10.1080/01434632.2020.1835929>
- Kutlu, E., Tiv, M., Wulff, S., & Titone, D. (2022). Does race impact speech perception? An account of accented speech in two different multilingual locales. *Cognitive Research: Principles and Implications*, 7(1), 1–16. <https://doi.org/10.1186/s41235-022-00354-0>

- Kymlicka, W. (2009). The current state of multiculturalism in Canada. *Canadian Journal for Social Research*, 2, 13–34.
- Lalonde, M. (2021, Oct. 15). *Holness says he receives racist and threatening messages 'non-stop'*. Vancouver Sun. <https://vancouver.sun.com/news/local-news/holness-says-he-receives-racist-and-threatening-messages-non-stop>
- Lambert, W. E., Frankle, H., & Tucker, G. R. (1966). Judging personality through speech: A French-Canadian example. *Journal of Communication*, 16, 305–321. <https://doi.org/10.1111/j.1460-2466.1966.tb00044.x>
- Lambert, W. E., Hodgson, R. C., Gardner, R. C., & Fillenbaum, S. (1960). Evaluational reactions to spoken languages. *The Journal of Abnormal and Social Psychology*, 60(1), 44–51. <https://doi.org/10.1037/h0044430>
- Lancee, B., & Dronkers, J. (2011). Ethnic, religious and economic diversity in Dutch neighbourhoods: Explaining quality of contact with neighbours, trust in the neighbourhood and inter-ethnic trust. *Journal of Ethnic and Migration Studies*, 37, 597–618. <https://doi.org/10.1080/1369183x.2011.545277>
- Lawlor, A., & Tolley, E. (2017). Deciding who's legitimate: News media framing of immigrants and refugees. *International Journal of Communication*, 11, 967–991.
- Lee, Y. Y., & Lin, J. (2005). Linking patients' trust in physicians to health outcomes. *British Journal of Hospital Medicine*, 69(1), 42–46. <https://doi.org/10.12968/hmed.2008.69.1.28040>
- Lee, B. J., & Bailey, J. L. (2022). Assumptions of speaker ethnicity and the effect on ratings of accentedness, comprehensibility, and intelligibility. *Language Awareness*, 1–22. <https://doi.org/10.1080/09658416.2022.2091143>

- Lenard, P. T. (2018). Wither the Canadian model? Evaluating the new Canadian nationalism (2006–2015). In J. E. Fossum, R. Kastoryano, & B. Siim (Eds.), *Diversity and contestations over nationalism in Europe and Canada* (pp. 211–236). Palgrave Macmillan.
https://doi.org/10.1057/978-1-137-58987-3_8
- Lev-Ari, S. (2018). Social network size can influence linguistic malleability and the Propagation of linguistic change. *Cognition*, *176*, 31–39. <https://doi.org/10.1016/j.cognition.2018.03.003>
- Lev-Ari, S., & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology*, *46*, 1093–1096.
<https://doi.org/10.1016/j.jesp.2010.05.025>
- Levine, M. (1988). The “reconquest” of Montreal: Public policy, language, and economic change 1960–1987. *Quebec Studies*, *6*, 41–64. <https://doi.org/10.3828/qs.6.1.41>
- Levinson, J. D., Smith, R. J., & Hioki, K. (2019). Race and retribution: An empirical study of implicit bias and punishment in America. *UC Davis Law Review*, *53*, 839–891.
- Lewis, K., Winsett, R. P., Cetingok, M., Martin, J., & Hathaway, D. K. (2000). Social network mapping with transplant recipients. *Progress in Transplantation*, *10*, 262–266.
<https://doi.org/10.7182/prtr.10.4.e455k83x2n7j7615>
- Linck, J. A., & Cunnings, I. (2015). The utility and application of mixed-effects models in second language research. *Language Learning*, *65*, 185–207. <https://doi.org/10.1111/lang.12117>
- Lindemann, S. (2003). Koreans, Chinese or Indians? Attitudes and ideologies about non-native English speakers in the United States. *Journal of Sociolinguistics*, *7*, 348–364.
<https://doi.org/10.1111/1467-9481.00228>
- Lindsay, C. (2007). *Profiles of ethnic communities in Canada*. Statistics Canada.

- Lippi-Green, R. (2012). English with an accent: Language, ideology, and discrimination in the United States. Routledge. <https://doi.org/10.4324/9780203348802>
- Lippi-Green, R. (1994). Accent, standard language ideology, and discriminatory pretext in the courts. *Language in Society*, 23, 163–198. <https://doi.org/10.1017/S0047404500017826>
- Locke, K. D. (2003). Status and solidarity in social comparison: agentic and communal values and vertical and horizontal directions. *Journal of Personality and Social Psychology*, 84, 619–631. <https://doi.org/10.1037/0022-3514.84.3.619>
- Lofaro, J., & Lurie, R. (2021, June 16). *Racially motivated hate crimes up by 53 per cent in Montreal last year*. CTV News. <https://montreal.ctvnews.ca/racially-motivated-hate-crimes-up-by-53-per-cent-in-montreal-last-year-report-1.5472848>
- Lou, N. M., & Noels, K. A. (2020). Mindsets matter for linguistic minority students: Growth mindsets foster greater perceived proficiency, especially for newcomers. *The Modern Language Journal*, 104, 739–756. <https://doi.org/10.1111/modl.12669>
- Louis, W. R., Esses, V. M., & Lalonde, R. N. (2013). National identification, perceived threat, and dehumanization as antecedents of negative attitudes toward immigrants in Australia and Canada. *Journal of Applied Social Psychology*, 43, E156–E165. <https://doi.org/10.1111/jasp.12044>
- Lu, Y., & Gnevshcheva, K. (2021). Accentedness and personality evaluation of Asian and Caucasian second language speakers of English by Asian second language English listeners. *Journal of Multilingual and Multicultural Development*, 1–12. <https://doi.org/10.1080/01434632.2021.1959598>
- Lund, D. E., & Hira-Friesen, P. (2014). Welcoming capacities of rural and urban Canadian communities. *Immigrant Integration: Research Implications for Future Policy*, 103–118. <https://doi.org/10.1353/ces.2013.0043>

- Lybaert, C., Van Hoof, S., & Deygers, B. (2022). The influence of ethnicity and language variation on undergraduates' evaluations of Dutch-speaking instructors in Belgium: A contextualized speaker evaluation experiment. *Language & Communication, 84*, 1–19.
<https://doi.org/10.1016/j.langcom.2022.01.004>
- Lybeck, K. (2002). *The role of acculturation and social networks in the acquisition of second language pronunciation*. [Doctoral dissertation]. University of Minnesota.
- Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods, 47*, 1122–1135. <https://doi.org/10.3758/s13428-014-0532-5>
- MacKinnon, P. (2022). *Canada in question: Exploring our citizenship in the twenty-first century*. University of Toronto Press. <https://doi.org/10.3138/9781487543167>
- Macvicar, A. (2019, September 30). 'Discrimination has no jurisdiction': Calgary council passes motion opposing Quebec's Bill 21. Global News.
<https://globalnews.ca/news/5972723/discrimination-has-no-jurisdiction-calgary-council-passes-motion-opposing-quebecs-bill-21/>
- Matsuda, M. J. (1991). Voices of America: Accent, antidiscrimination law, and a jurisprudence for the last reconstruction. *Yale Law Journal, 100*, 1329–1407. <https://doi.org/10.2307/796694>
- Mbakogu, I., Duhaney, P., Ferrer, I., & Lee, E. (2021). Confronting whiteness in social work education through racialized student activism. *Canadian Social Work Review, 38*(2), 113–140.
<https://doi.org/10.7202/1086122ar>
- McDaniel, E. L., Nooruddin, I., & Faith Shortle, A. (2011). Divine boundaries: How religion shapes citizens' attitudes toward immigrants. *American Politics Research, 39*(1), 205–233.
<https://doi.org/10.1177/1532673X10371300>

- McDonald, D. (2014). Credibility assessment in refugee status determination. *National Law School of India Review*, 26(2), 115–126.
- McGowan, K. B. (2015). Social expectation improves speech perception in noise. *Language and Speech*, 58, 502–521. <https://doi.org/10.1177/0023830914565191>
- McKenzie, R. M. (2015). UK university students' folk perceptions of spoken variation in English: The role of explicit and implicit attitudes. *International Journal of the Sociology of Language*, 2015(236), 31–53. <https://doi.org/10.1515/ijsl-2015-0020>
- McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27, 415–444. <https://doi.org/10.1146/annurev.soc.27.1.415>
- Mepham, K. D., & Martinovic, B. (2018). Multilingualism and out-group acceptance: The mediating roles of cognitive flexibility and deprovincialization. *Journal of Language and Social Psychology*, 37(1), 51–73. <https://doi.org/10.1177/0261927x17706944>
- Miconi, D., Li, Z. Y., Frounfelker, R. L., Venkatesh, V., & Rousseau, C. (2021). Socio-cultural correlates of self-reported experiences of discrimination related to COVID-19 in a culturally diverse sample of Canadian adults. *International Journal of Intercultural Relations*, 81, 176–192. <https://doi.org/10.1016/j.ijintrel.2021.01.013>
- Milton, P. (2006). Multilingualism: The Canadian way. *Education Canada*, 46(4), 55–56.
- Mirshahidi, S. (2017). I find you attractive but I don't trust you: The case of language attitudes in Iran. *Journal of Multilingual and Multicultural Development*, 38(2), 146–159. <https://doi.org/10.1080/01434632.2016.1178268>

- Monaghan, J., & Santos, M. (2020). Canada, the Palestinian Liberation Organisation (PLO) and the flexibility of terror identities. *Critical Studies on Terrorism, 13*, 280–295.
<https://doi.org/10.1080/17539153.2020.1722308>
- Montreuil, A., & Bourhis, R. Y. (2001). Majority acculturation orientations toward “valued” and “devalued” immigrants. *Journal of Cross-Cultural Psychology, 32*, 698–719.
<https://doi.org/10.1177/0022022101032006004>
- Montreuil, A., Bourhis, R. Y., & Vanbeselaere, N. (2004). Perceived threat and host community acculturation orientations towards immigrants: Comparing Flemings in Belgium and Francophones in Québec. *Canadian Ethnic Studies, 36*(3), 113–135.
- Munro, M. J. (2003). A primer on accent discrimination in the Canadian context. *TESL Canada Journal, 20*(2), 38–51. <https://doi.org/10.18806/tesl.v20i2.947>
- Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 45*(1), 73–97.
<https://doi.org/10.1111/0023-8333.49.s1.8>
- Nadal, K. L., Davidoff, K. C., Davis, L. S., Wong, Y., Marshall, D., & McKenzie, V. (2015). A qualitative approach to intersectional microaggressions: Understanding influences of race, ethnicity, gender, sexuality, and religion. *Qualitative Psychology, 2*(2), 147–163.
<https://doi.org/10.1037/qup0000026>
- Nadal, K. L., Wong, Y., Griffin, K. E., Davidoff, K., & Sriken, J. (2014). The adverse impact of racial microaggressions on college students' self-esteem. *Journal of College Student Development, 55*, 461–474. <https://doi.org/10.1353/csd.2014.0051>

- Nagle, C. (2019). Developing and validating a methodology for crowdsourcing L2 speech ratings in Amazon Mechanical Turk. *Journal of Second Language Pronunciation*, 5, 294–323.
<https://doi.org/10.1075/jslp.18016.nag>
- Nagle, C. L., & Rehman, I. (2021). Doing L2 speech research online: Why and how to collect online ratings data. *Studies in Second Language Acquisition*, 43, 916–939.
<https://doi.org/10.1017/s0272263121000292>
- Nangia, P. (2013). *Discrimination experienced by landed immigrants in Canada*. Ryerson Centre for Immigration and Settlement. <https://doi.org/10.32920/ryerson.14639880.v2>
- Narayan, M. C. (2019). CE: Addressing implicit bias in nursing: A review. *The American Journal of Nursing*, 119(7), 36–43. <https://doi.org/10.1097/01.naj.0000569340.27659.5a>
- Nemțoi, G., & Ignătescu, C. (2014). National citizenship as representative of European citizenship. *Procedia-Social and Behavioral Sciences*, 149, 653–658.
<https://doi.org/10.1016/j.sbspro.2014.08.243>
- Neuman, K. (2019). *Race relations in Canada 2019: A survey of Canadian public opinion and experience*. Canadian Race Relations Foundation.
<https://www.environicsinstitute.org/projects/project-details/race-relations-in-canada-2019>
- Ng, S. H. (2007). Language-based discrimination: Blatant and subtle forms. *Journal of Language and Social Psychology*, 26(2), 106–122. <https://doi.org/10.1177/0261927x07300074>
- Nghia, T. L. H., Singh, J. K. N., Pham, T., & Medica, K. (2020). Employability, employability capital, and career development. In T. L. H. Nghia, T. Pham, M. Tomlinson, K. Medica, & C. Thompson (Eds.). *Developing and utilizing employability capitals: Graduates' strategies across labour markets*. Routledge. <https://doi.org/10.4324/9781003004660-4>

- Nguyen, B. B. D. (1993). Accent discrimination and the test of spoken English: A call for an objective assessment of the comprehensibility of nonnative speakers. *California Law Review*, *81*, 1325–1361. <https://doi.org/10.2307/3480920>
- Nilsson, N. J. (2014). *Understanding beliefs*. MIT Press.
<https://doi.org/10.7551/mitpress/10055.001.0001>
- Oppong, E. (2018). Health care choices of Ghanaian adult immigrants in Calgary, Alberta, Canada. *European Journal of Public Health*, *28*, 44–67. <https://doi.org/10.1093/eurpub/cky047.074>
- Oxman-Martinez, J., Rummens, A. J., Moreau, J., Choi, Y. R., Beiser, M., Ogilvie, L., & Armstrong, R. (2012). Perceived ethnic discrimination and social exclusion: Newcomer immigrant children in Canada. *American Journal of Orthopsychiatry*, *82*, 376–388. <https://doi.org/10.1111/j.1939-0025.2012.01161.x>
- Palardy, C., & Filip, R. (2015). *Portraits statistiques: L'immigration temporaire au Québec, 2008–2013*. Quebec: Ministère de l'immigration, la diversité et l'inclusion.
- Pantos, A. J., & Perkins, A. W. (2013). Measuring implicit and explicit attitudes toward foreign accented speech. *Journal of Language and Social Psychology*, *32*(1), 3–20.
<https://doi.org/10.1177/0261927x12463005>
- Pauker, K., Carpinella, C., Meyers, C., Young, D. M., & Sanchez, D. T. (2018). The role of diversity exposure in Whites' reduction in race essentialism over time. *Social Psychological and Personality Science*, *9*, 944–952. <https://doi.org/10.1177/1948550617731496>
- Pendakur, K., & Pendakur, R. (2002). Language as both human capital and ethnicity. *International Migration Review*, *36*(1), 147–177. <https://doi.org/10.1111/j.1747-7379.2002.tb00075.x>
- Pettigrew, T. F. (1998). Intergroup contact theory. *Annual Review of Psychology*, *49*(1), 65–85.
<https://doi.org/10.1146/annurev.psych.49.1.65>

- Pietraszewski, D., & Schwartz, A. (2014). Evidence that accent is a dedicated dimension of social categorization, not a byproduct of coalitional categorization. *Evolution and Human Behavior*, 35(1), 51–57. <https://doi.org/10.1016/j.evolhumbehav.2013.09.005>
- Piller, I. (2001). Naturalization language testing and its basis in ideologies of national identity and citizenship. *International Journal of Bilingualism*, 5, 259–277. <https://doi.org/10.1177/13670069010050030201>
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64, 878–912. <https://doi.org/10.1111/lang.12079>
- Portes, A., & Zhou, M. (1993). The new second generation: Segmented assimilation and its variants. *Annals of the American Academy of Political and Social Science*, 530(1), 74–96. <https://doi.org/10.4324/9780429499821-51>
- Preisler, S. (2021). *The vulnerability of newcomers (including refugees) in the Canadian housing sector*. Community Housing Canada. https://era.library.ualberta.ca/items/6b99aeb6-9f6c-4caa-b998-b6416d09b63e/view/9d721a6b-14ca-4a1f-8473-10ac8c0c0c9d/Newcomers_Final.pdf
- Putnam, R. (1993). The prosperous community: Social capital and public life. *The American Prospect*, 13 (4), 1–11.
- R Core Team. (2020). *R: A language and environment for statistical computing*. <https://www.R-project.org/>
- Rahmath, S., Chambers, L., & Wakewich, P. (2016, September). Asserting citizenship: Muslim women’s experiences with the hijab in Canada. *Women’s Studies International Forum*, 58, 34–40. <https://doi.org/10.1016/j.wsif.2016.06.001>

- Rakić, T., Steffens, M. C., & Mummendey, A. (2011). When it matters how you pronounce it: The influence of regional accents on job interview outcome. *British Journal of Psychology*, *102*, 868–883. <https://doi.org/10.1111/j.2044-8295.2011.02051.x>
- Raney, T. (2009). As Canadian as possible... under what circumstances? Public opinion on national identity in Canada outside Quebec. *Journal of Canadian Studies*, *43*(3), 5–29. <https://doi.org/10.3138/jcs.43.3.5>
- Rapoport, A. (2021). The changing meaning of citizenship and identity and a perspective model of citizenship education. In *Globalisation, cultural identity and nation-building* (pp. 37–50). Springer, Dordrecht. https://doi.org/10.1007/978-94-024-2014-2_3
- Razack, S. H. (2007). “Your client has a profile:” Race and national security in Canada after 9/11. In *Studies in Law, Politics and Society*. Emerald Group Publishing Limited. [https://doi.org/10.1016/s1059-4337\(06\)40001-6](https://doi.org/10.1016/s1059-4337(06)40001-6)
- Rebhun, U. (2021). Immigrant integration and COVID-19. *Border Crossing*, *11*(1), 17–23. <https://doi.org/10.33182/bc.v11i1.1291>
- Risager, K. (2006). *Language and culture*. Multilingual Matters. <https://doi.org/10.21832/9781853598609>
- Romaniuk, A. (2017). Stationary population, immigration, social cohesion, and national identity: What are the links and the policy implications? With special attention to Canada, a demographer’s point of view. *Canadian Studies in Population*, *44*, 165–178. <https://doi.org/10.15407/dse2018.01.011>
- Rouhana, N. N., & Fiske, S. T. (1995). Perception of power, threat, and conflict intensity in asymmetric intergroup conflict: Arab and Jewish citizens of Israel. *Journal of Conflict Resolution*, *39*(1), 49–81. <https://doi.org/10.1177/0022002795039001003>

- Rubin, D. L. (1992). Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. *Research in Higher Education*, 33, 511–531.
<https://doi.org/10.1007/BF00973770>
- Rubin, D. L., & Smith, K. A. (1990). Effects of accent, ethnicity, and lecture topic on undergraduates' perceptions of nonnative English-speaking teaching assistants. *International Journal of Intercultural Relations*, 14, 337–353. [https://doi.org/10.1016/0147-1767\(90\)90019-s](https://doi.org/10.1016/0147-1767(90)90019-s)
- Rushton, J. P., & Bons, T. A. (2005). Mate choice and friendship in twins: evidence for genetic similarity. *Psychological Science*, 16, 555–559. <https://doi.org/10.1111/j.0956-7976.2005.01574.x>
- Ryan, E. B., Carranza, M. A., & Moffie, R. W. (1977). Reactions toward varying degrees of accentedness in the speech of Spanish-English bilinguals. *Language and Speech*, 20, 267–273.
<https://doi.org/10.1177/002383097702000308>
- Sa'd, S. H. T. (2018). Learners' views of (non) native speaker status, accent, and identity: an English as an international language perspective. *Journal of World Languages*, 5(1), 1–22.
<https://doi.org/10.1080/21698252.2018.1500150>
- Saito, K., Trofimovich, P., & Isaacs, T. (2017). Using listener judgments to investigate linguistic influences on L2 comprehensibility and accentedness: A validation and generalization study. *Applied Linguistics*, 38, 439–462. <https://doi.org/10.1093/applin/amv047>
- Saito, K., Trofimovich, P., & Isaacs, T. (2016). Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. *Applied Psycholinguistics*, 37, 217–240. <https://doi.org/10.1017/s0142716414000502>

- Sam, D. L., & Berry, J. W. (1995). Acculturative stress among young immigrants in Norway. *Scandinavian Journal of Psychology*, 36(1), 10–24. <https://doi.org/10.1111/j.1467-9450.1995.tb00964.x>
- Saribay, S. A., Biten, A. F., Meral, E. O., Aldan, P., Třebický, V., & Kleisner, K. (2018). The Bogazici face database: Standardized photographs of Turkish faces with supporting materials. *PLoS One*, 13, e0192018. <https://doi.org/10.1371/journal.pone.0192018>
- Scheepers, P., Gijssberts, M., & Coenders, M. (2002). Ethnic exclusionism in European countries. Public opposition to civil rights for legal migrants as a response to perceived ethnic threat. *European Sociological Review*, 18(1), 17–34. <https://doi.org/10.1093/esr/18.1.17>
- Schiffer, E., & Hauck, J. (2010). Net-Map: Collecting social network data and facilitating network learning through participatory influence network mapping. *Field Methods*, 22, 231–249. <https://doi.org/10.1177/1525822X10374798>
- Schissel, B., Wanner, R., & Frideres, J. S. (1989). Social and economic context and attitudes toward immigrants in Canadian cities. *International Migration Review*, 23, 289–308. <https://doi.org/10.2307/2546262>
- Schluter, A. A. (2021). Atatürk's long shadow: standard Turkish speakers as younger, more successful, and more attractive than their Kurdish-accented regional counterparts. *Journal of Multilingual and Multicultural Development*, 42, 840–853. <https://doi.org/10.1080/01434632.2020.1822851>
- Schmidt, C., Young, J., & Mandzuk, D. (2010). The integration of immigrant teachers in Manitoba, Canada: Critical issues and perspectives. *Journal of International Migration and Integration*, 11, 439–452. <https://doi.org/10.1007/s12134-010-0149-1>
- Scott, A. (2020). *Creating a welcoming community: A toolkit to support immigrants, refugees, and BIPOC*. University of Minnesota Extension.

- Sharples, A. E., & Chasteen, A. L. (2021). Unpacking the complexity of immigrant attitudes in Canada: Immigrant age predicts attitudes toward immigrants. *Canadian Journal of Behavioural Science*, 53(1), 23–35. <https://doi.org/10.1037/cbs0000166>
- Sherif, M. (1966). *In common predicament: Social psychology of intergroup conflict and cooperation*. Houghton Mifflin. <https://doi.org/10.2307/2091843>
- Siegel, J. (2003). Social context. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 178–223). Blackwell. <https://doi.org/10.1002/9780470756492.ch8>
- Silverstein, M. (2015). How language communities intersect: Is “superdiversity” an incremental or transformative condition?. *Language & Communication*, 44, 7–18. <https://doi.org/10.1016/j.langcom.2014.10.015>
- Simon, E., Lybaert, C., & Plevoets, K. (2022). Social attitudes, intelligibility and comprehensibility: The role of the listener in the perception of non-native speech. *Vigo International Journal of Applied Linguistics*, 19, 177–222. <https://doi.org/10.35869/vial.v0i19.3763>
- Simonsen, K. B. (2016). How the host nation's boundary drawing affects immigrants’ belonging. *Journal of Ethnic and Migration Studies*, 42, 1153–1176. <https://doi.org/10.1080/1369183x.2016.1138854>
- Singer, C. (2019). *Punjabi now third language in Parliament of Canada*. Immigration Canada. <https://www.immigration.ca/punjabi-now-third-language-in-parliament-of-canada>
- Singh, L., Quinn, P. C., Qian, M., & Lee, K. (2020). Bilingualism is associated with less racial bias in preschool children. *Developmental Psychology*, 56, 888–896. <https://doi.org/10.1037/dev0000905>

- Sioufi, R., Bourhis, R. Y., & Allard, R. (2015). Vitality and ethnolinguistic attitudes of Acadians, Franco-Ontarians and Francophone Quebecers: Two or three solitudes in Canada's bilingual belt?. *Journal of Multilingual and Multicultural Development*, 37, 385–401.
<https://doi.org/10.1080/01434632.2015.1072205>
- Smith, A. (2018). Mirror image: The fight against homelessness in Calgary and Montreal. In S. Breux & J. Couture (Eds.), *Accountability and responsiveness at the municipal level: Views from Canada* (pp. 223–240). McGill-Queen's Press. <https://doi.org/10.2307/j.ctv2n7pbt.14>
- Smith, G. B. (2005). I want to speak like a native speaker: The case for lowering the plaintiff's burden of proof in Title VII accent discrimination cases. *Ohio State Law Journal*, 66, 231–267.
- Sokolov, C. (2014). Self-evaluation of rater bias in written composition assessment. *Linguistica*, 54, 261–275. <https://doi.org/10.4312/linguistica.54.1.261-275>
- Soroka, S. N., & Roberton, S. (2010). *A literature review of Public Opinion Research on Canadian attitudes towards multiculturalism and immigration, 2006–2009*. Citizenship and Immigration Canada, Research and Evaluation.
<https://www.canada.ca/content/dam/ircc/migration/ircc/english/pdf/research-stats/2012-por-multi-imm-eng.pdf>
- Sparrowe, R.T., Liden, R.C., Wayne, S.J., & Kraimer, M.L. (2001). Social networks and performance of individuals and groups. *Academy of Management Journal*, 44, 318–325.
<http://dx.doi.org/10.2307/3069458>
- Statistics Canada. (2017). *Census profile*. <https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/index.cfm?Lang=E>

- Stephan, W. G., and Stephan, C. W. (2000). An integrated threat theory of prejudice. In S. Oskamps (Ed.), *Reducing prejudice and discrimination: The Claremount Symposium on Applied Social Psychology* (pp. 23–45). Lawrence Erlbaum. <https://doi.org/10.4324/9781410605634-7>
- Stocker, L. (2017). The impact of foreign accent on credibility: An analysis of cognitive statement ratings in a Swiss Context. *Journal of Psycholinguistic Research*, *46*, 617–628. <https://doi-org.lib-ezproxy.concordia.ca/10.1007/s10936-016-9455-x>
- Strauss, D. J., & Francis, A. L. (2017). Toward a taxonomic model of attention in effortful listening. *Cognitive, Affective, & Behavioral Neuroscience*, *17*, 809–825. <https://doi.org/10.3758/s13415-017-0513-0>
- Sumantry, D., & Choma, B. L. (2021). Accent-based stereotyping, prejudice, and their predictors. *Personality and Individual Differences*, *179*, 1–6. <https://doi.org/10.1016/j.paid.2021.110894>
- Sumner, M. (2011). The role of variation in the perception of accented speech. *Cognition*, *119*(1), 131–136. <https://doi.org/10.1016/j.cognition.2010.10.018>
- Suraweera, D. (2019). *Language discrimination at English dominant workplaces* [Paper presentation]. Canadian Association for the Study of Adult Education, Montreal, QC, Canada.
- Swain, M., & Lapkin, S. (2005). The evolving sociopolitical context of immersion education in Canada: some implications for program development 1. *International Journal of Applied Linguistics*, *15*(2), 169–186. <https://doi.org/10.1111/j.1473-4192.2005.00086.x>
- Tajfel, H. (1974). Social identity and intergroup behaviour. *International Social Science Council*, *13*(2), 65–93. <https://doi.org/10.1177/053901847401300204>
- Tajfel, H. (1972). Experiments in a vacuum. In J. Israel & H. T. Triandis (Eds.), *The context of social psychology: A critical assessment* (pp. 69–119). Academic Press.

- Tajfel, H. (1959). A note on Lambert's 'Evaluational reactions to spoken languages.' *Canadian Journal of Psychology*, *13*(2), 86–92. <https://doi.org/10.1037/h0083762>
- Tan, J. H., & Vogel, C. (2008). Religion and trust: An experimental study. *Journal of Economic Psychology*, *29*, 832–848. <https://doi.org/10.1016/j.joep.2008.03.002>
- Taylor Reid, K., Trofimovich, P. & O'Brien, M. G. (2019). Social attitudes and speech ratings: Effects of positive and negative bias on multiage listeners' judgments of second language speech. *Studies in Second Language Acquisition*, *41*, 419–442. <https://doi.org/10.1017/S0272263118000244>
- Timming, A. R. (2017). The effect of foreign accent on employability: A study of the aural dimensions of aesthetic labour in customer-facing and non-customer-facing jobs. *Work, Employment and Society*, *31*, 409–428. <https://doi.org/10.1177/0950017016630260>
- Tracy, E. M., & Whittaker, J. K. (1990). The social network map: Assessing social support in clinical practice. *Families in Society*, *71*, 461–470. <https://doi.org/10.1177/104438949007100802>
- Tropp, L. R., & Pettigrew, T. F. (2005). Relationships between intergroup contact and prejudice among minority and majority status groups. *Psychological Science*, *16*, 951–957. <https://doi.org/10.1111/j.1467-9280.2005.01643.x>
- Uther, M., Knoll, M. A., & Burnham, D. (2007). Do you speak E-NG-LI-SH? A comparison of foreigner-and infant-directed speech. *Speech Communication*, *49*(1), 2–7. <https://doi.org/10.1016/j.specom.2006.10.003>
- Valentino, N. A., Brader, T., & Jardina, A. E. (2013). Immigration opposition among US Whites: General ethnocentrism or media priming of attitudes about Latinos?. *Political Psychology*, *34*(2), 149–166. <https://doi.org/10.1111/j.1467-9221.2012.00928.x>

- Vang, Z. M., Hou, F., & Elder, K. (2019). Perceived religious discrimination, religiosity, and life satisfaction. *Journal of Happiness Studies*, 20, 1913–1932. <https://doi.org/10.1007/s10902-018-0032-x>
- Veenstra, G. (2011). Mismatched racial identities, colourism, and health in Toronto and Vancouver. *Social Science & Medicine*, 73, 1152–1162. <https://doi.org/10.1016/j.socscimed.2011.07.030>
- Verkuyten, M., & Thijs, J. (2013). Multicultural education and inter-ethnic attitudes: An intergroup perspective. *European Psychologist*, 18(3), 179–190. <https://doi.org/10.1027/1016-9040/a000152>
- Vermeij, L., Van Duijn, M. A., & Baerveldt, C. (2009). Ethnic segregation in context: Social discrimination among native Dutch pupils and their ethnic minority classmates. *Social Networks*, 31, 230–239. <https://doi.org/10.1016/j.socnet.2009.06.002>
- Vickers, L. (2004). Approaching religious discrimination at work: Lessons from Canada. *International Journal of Comparative Labour Law and Industrial Relations*, 20(2), 177–200. <https://doi.org/10.54648/ijcl2004011>
- Viola, S. (2012). Intercultural welcoming spaces in Montréal: Harmonization drivers for a new sense of identity. *City, Culture and Society*, 3(2), 141–149. <https://doi.org/10.1016/j.ccs.2012.01.001>
- Vomiero, J. (2018, May 16). *If immigration was slashed, Canada's economy would feel the pinch*. Global News. <https://globalnews.ca/news/4211243/immigration-canadian-economy/>
- Von Neumann, J. & Morgenstern, O. (1944). *The theory of games and economic behavior*. Princeton University Press. <https://doi.org/10.1515/9781400829460>
- Walls, M. L., Gonzalez, J., Gladney, T., & Onello, E. (2015). Unconscious biases: Racial microaggressions in American Indian health care. *The Journal of the American Board of Family Medicine*, 28, 231–239. <https://doi.org/10.3122/jabfm.2015.02.140194>

- Weichselbaumer, D. (2020). Multiple discrimination against female immigrants wearing headscarves. *ILR Review*, 73, 600–627. <https://doi.org/10.1177/0019793919875707>
- Wheeler, M., de Bourmont, S., Paul-Emile, K., Pfeffinger, A., McMullen, A., Critchfield, J. M., & Fernandez, A. (2019). Physician and trainee experiences with patient bias. *JAMA Internal Medicine*, 179, 1678–1685. <https://doi.org/10.1001/jamainternmed.2019.4122>
- White, R. (2021 Aug. 4). *Calgary police seek man who hurled racial slurs at woman in downtown coffee dispute*. CTV News. <https://calgary.ctvnews.ca/calgary-police-seek-man-who-hurled-racial-slurs-at-woman-in-downtown-coffee-dispute-1.5559427>
- Wilkins-Laflamme, S. (2018). Islamophobia in Canada: Measuring the realities of negative attitudes toward Muslims and religious discrimination. *Canadian Review of Sociology*, 55(1), 86–110. <https://doi.org/10.1111/cars.12180>
- Williams, M. T., Khanna Roy, A., MacIntyre, M. P., & Faber, S. (2022). The traumatizing impact of racism in Canadians of colour. *Current Trauma Reports*, 8, 17–34. <https://doi.org/10.1007/s40719-022-00225-5>
- Wilson, T. C. (2001). Americans' views on immigration policy: Testing the role of threatened group interests. *Sociological Perspectives*, 44, 485–501. <https://doi.org/10.2307/1389655>
- Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological review*, 107(1), 101–126. <https://doi.org/10.1037/0033-295X.107.1.101>
- Winke, P., Gass, S., & Myford, C. (2013). Raters' L2 background as a potential source of bias in rating oral performance. *Language Testing*, 30, 231–252. <https://doi.org/10.1177/0265532212456968>
- Winter, E. (2014). (Im) possible citizens: Canada's 'citizenship bonanza' and its boundaries. *Citizenship Studies*, 18(1), 46–62. <https://doi.org/10.1080/13621025.2012.707010>

- Wong, P., Lai, C. F., Nagasawa, R., & Lin, T. (1998). Asian Americans as a model minority: Self-perceptions and perceptions by other racial groups. *Sociological Perspectives, 41*(1), 95–118. <https://doi.org/10.2307/1389355>
- Wu, Z., & Finnsdottir, M. (2021). Perceived racial and cultural discrimination and sense of belonging in Canadian society. *Canadian Review of Sociology, 58*, 229–249. <https://doi.org/10.1111/cars.12339>
- Wu, Z., & Schimmele, C. M. (2021). Perceived religious discrimination and mental health. *Ethnicity & Health, 26*, 963–980. <https://doi.org/10.1080/13557858.2019.1620176>
- Wu, C., Wilkes, R., Qian, Y., & Kennedy, E. B. (2020). *Acute discrimination and East Asian-white mental health gap during COVID-19 in Canada*. (Working Paper SSRN 3626460).
- Yan, X., & Ginther, A. (2017). Listeners and raters: Similarities and differences in evaluation of accented speech. In O. Kang & A. Ginther (Eds.), *Assessment in second language pronunciation* (pp. 67–88). London: Routledge. <https://doi.org/10.4324/9781315170756-5>
- Yi, H. G., Phelps, J. E., Smiljanic, R., & Chandrasekaran, B. (2013). Reduced efficiency of audiovisual integration for nonnative speech. *The Journal of the Acoustical Society of America, 134*, EL387–EL393. <https://doi.org/10.1121/1.4822320>
- Zheng, Y., & Samuel, A. G. (2017). Does seeing an Asian face make speech sound more accented?. *Attention, Perception, & Psychophysics, 79*, 1841–1859. <https://doi.org/10.3758/s13414-017-1329-2>

Appendix A

Sample Image Guises for Speech Rating Task



Appendix B

Image Pre-Rating Task

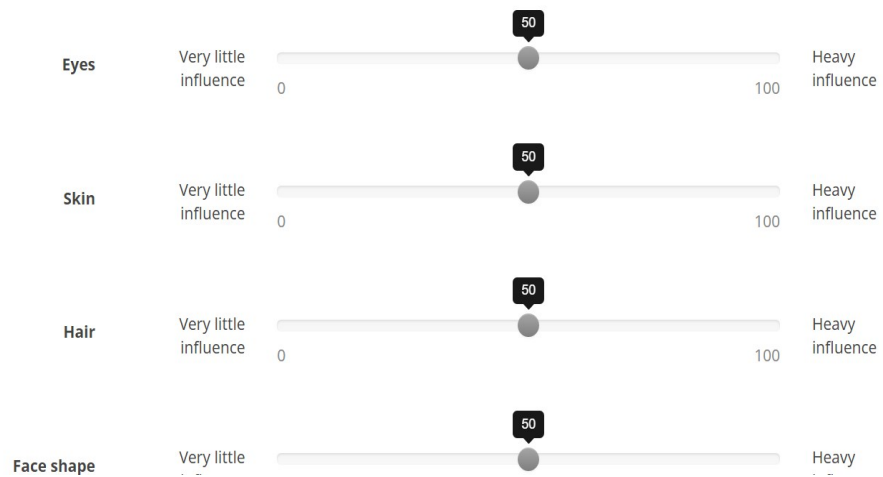
*



Of which primary ethnic origin does this person appear to be?

- South Asian (Nepalese, Indian, Bangladeshi, Pakistani, etc.)
- Middle Eastern (Iraqi, Syrian, Turkish, Saudi Arabian, Lebanese, etc.)
- East Asian (Chinese, Japanese, Korean, Mongolian, etc.)
- European (Dutch, Irish, Norwegian, Polish, etc.)
- Southeast Asian (Cambodian, Filipino, Indonesian, Thai, etc.)
- Other:

*To what extent did the following features influence your choice?



Are there any other features that helped you to decide on this ethnic origin? If so, please explain.

*To what extent is this person's appearance typical for someone with the ethnic origin you chose?



Click on the circle and drag left or right to make your choice.

*How familiar are you with this ethnic origin?



Click on the circle and drag left or right to make your choice.

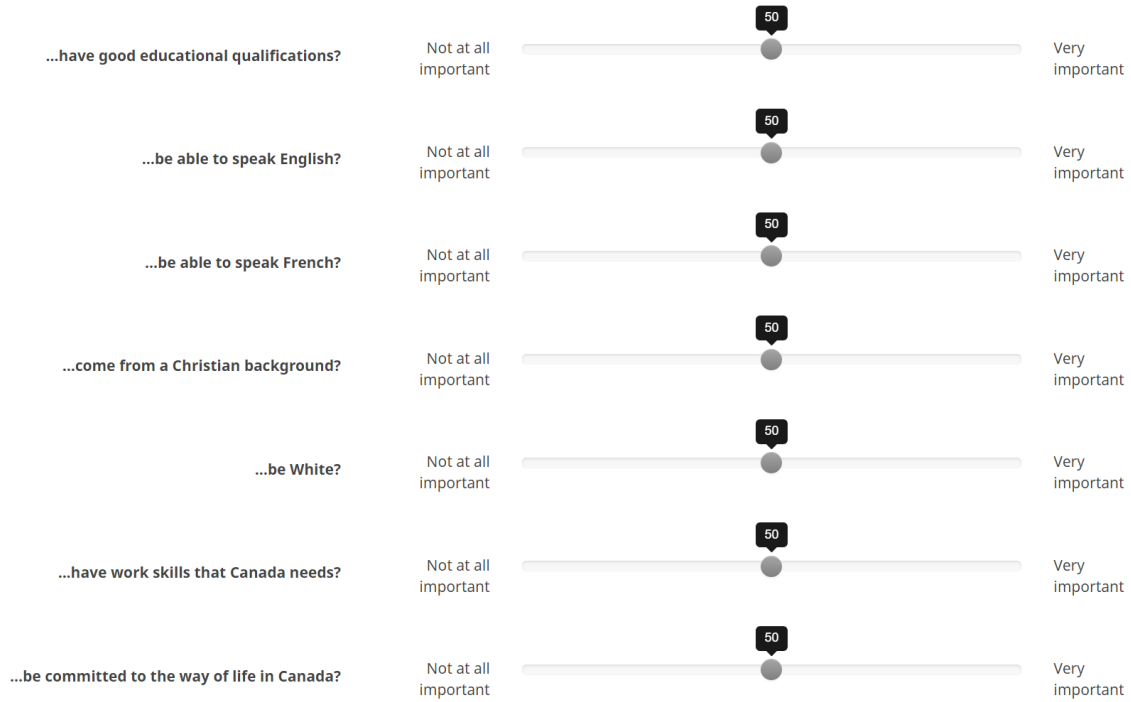
Appendix C

Social Attitudes Questionnaire

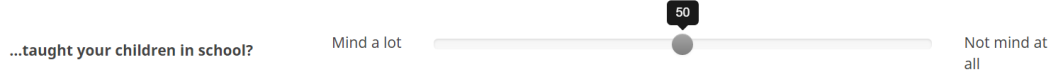
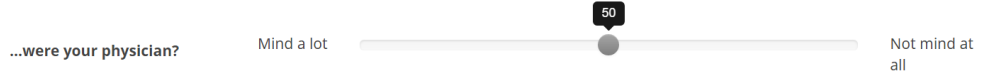
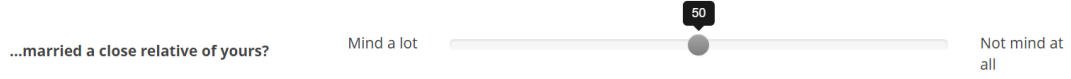
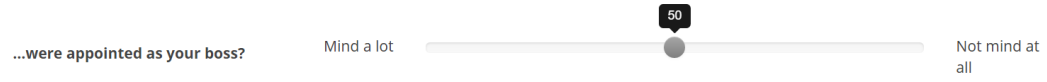
People come to live in Canada from other countries for various reasons. Some have ancestral ties. Others come to work here or join their families. Others come because they are under threat. Below are some questions about this issue.

*Please share how important you think each of the following factors should be in deciding whether someone born, raised, or living outside of Canada should be able to come live here.

How important should it be for them to...



*How much would you mind or not mind if someone from another country who is a *different race or ethnic group* from most Canadian people...



Appendix D

Full Statistical Output for Mixed-Effects Models in Study 1

Table D1

Summary of Mixed-Effects Model for Arabic Speakers' Comprehensibility Ratings in the Image and Voice Match and Mismatch Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	60.20	13.22	[29.43, 90.99]	4.55	.104
Voice-image match	2.51	4.41	[-6.13, 11.14]	0.57	.571
Voice-image mismatch	2.05	4.41	[-6.59, 10.69]	0.47	.643
Location	5.35	2.36	[0.72, 9.95]	2.27	.024
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	169.20	13.01	Log-likelihood	-1625.70	
Speaker gender (intercept)	314.10	17.72	AIC	3265.40	
			BIC	3292.60	
			Marginal R^2	0.01	
			Conditional R^2	0.61	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels. AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table D2

Summary of Mixed-Effects Model for Mandarin Speakers' Comprehensibility Ratings in the Image and Voice Match and Mismatch Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	73.19	5.05	[63.79, 82.58]	14.48	< .001
Voice-image match	-0.23	4.52	[-9.12, 8.66]	-0.05	.959
Voice-image mismatch	2.08	4.52	[-6.81, 10.97]	0.46	.646
Location	-3.48	2.32	[-8.02, 1.07]	-1.50	.135
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	233.29	15.27	Log-likelihood	-1614.10	
Speaker gender (intercept)	13.36	3.66	AIC	3242.20	
			BIC	3269.40	
			Marginal R^2	0.01	
			Conditional R^2	0.52	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels. AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table D3

Summary of Mixed-Effects Model for English Speakers' Comprehensibility Ratings in the Image and Voice Match and Mismatch Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	96.20	1.66	[92.98, 99.42]	58.11	< .001
Voice-image match	0.01	1.70	[-3.32, 3.34]	0.01	.997
Voice-image mismatch	0.79	1.70	[-2.54, 4.12]	0.46	.644
Location	0.35	0.83	[-1.28, 1.97]	0.42	.677
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	41.59	6.45	Log-likelihood	-1238.00	
Speaker gender (intercept)	0.08	0.28	AIC	2489.90	
			BIC	2517.20	
			Marginal R^2	0.01	
			Conditional R^2	0.70	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels. AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table D4

Summary of Mixed-Effects Model for Tagalog Speakers' Comprehensibility Ratings in the Image and Voice Match and Mismatch Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	83.66	8.33	[64.44, 102.88]	10.05	.035
Voice-image match	2.86	3.13	[-3.29, 9.01]	0.91	.364
Voice-image mismatch	0.80	3.13	[-5.35, 6.95]	0.26	.799
Location	0.09	1.65	[-3.14, 3.32]	0.05	.957
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	94.74	9.73	Log-likelihood	-1496.60	
Speaker gender (intercept)	120.61	10.98	AIC	3007.20	
			BIC	3034.40	
			Marginal R^2	0.01	
			Conditional R^2	0.60	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels. AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table D5

Summary of Mixed-Effects Model for Punjabi Speakers' Comprehensibility Ratings in the Image and Voice Match and Mismatch Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	61.84	4.78	[52.50, 71.18]	12.94	< .001
Voice-image match	1.69	4.98	[-8.05, 11.42]	0.34	.735
Voice-image mismatch	0.05	4.98	[-9.68, 9.78]	0.01	.992
Location	0.47	2.58	[-4.56, 5.51]	0.18	.854
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	271.00	16.46	Log-likelihood	-1650.40	
Speaker gender (intercept)	0.00	0.00	AIC	3314.80	
			BIC	3342.00	
			Marginal R^2	0.01	
			Conditional R^2	0.47	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels. AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table D6*Summary of Mixed-Effects Model for Combined Comprehensibility Ratings*

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	64.66	3.52	[56.41, 72.90]	18.38	.016
Mandarin speakers	7.37	1.25	[4.92, 9.83]	5.88	< .001
English speakers	31.83	1.25	[29.37, 34.28]	25.39	< .001
Punjabi speakers	-2.05	1.25	[-4.51, 0.40]	-1.64	.102
Tagalog speakers	20.43	1.25	[17.97, 22.88]	16.30	< .001
Location	0.59	1.41	[-2.17, 3.35]	0.42	.676
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	113.80	10.67	Log-likelihood	-7827.60	
Speaker gender (intercept)	21.51	4.64	AIC	15673.00	
			BIC	15723.00	
			Marginal R^2	0.28	
			Conditional R^2	0.52	

Note. The Arabic speaker group and Calgary were set as reference levels. AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table D7

Estimated Marginal Means of Comprehensibility Ratings for Speaker Groups and Between-Group Comparisons

Speaker group	EMM (1.08)	Speaker group	EMM _{diff} (1.28)	95% CI
Arabic	64.95	Mandarin	-7.37	[-11.13, -3.61]
		English	-31.83	[-35.59, -28.07]
		Punjabi	2.05	[-1.71, 5.81]
		Tagalog	-20.43	[-24.19, -16.67]
Mandarin	72.32	English	-24.46	[-28.22, -20.69]
		Punjabi	9.43	[5.66, 13.19]
		Tagalog	-13.06	[-16.82, -9.29]
English	96.78	Punjabi	33.88	[30.12, 37.64]
		Tagalog	11.40	[7.64, 15.16]
Punjabi	62.90	Tagalog	-22.48	[-26.24, -18.72]
Tagalog	85.38			

Table D8

Summary of Mixed-Effects Model for Arabic Speakers' Accentedness Ratings in the Image and Voice Match and Mismatch Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	76.05	7.15	[60.33, 91.76]	10.64	.013
Voice-image match	-0.73	3.77	[-8.10, 6.64]	-0.19	.846
Voice-image mismatch	-3.31	3.77	[-10.68, 4.07]	-0.88	.381
Location	-0.95	1.84	[-4.54, 2.65]	-0.52	.606
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	205.66	14.34	Log-likelihood	-1526.10	
Speaker gender (intercept)	75.73	8.70	AIC	3066.20	
			BIC	3093.40	
			Marginal R^2	0.015	
			Conditional R^2	0.77	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels. AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table D9

Summary of Mixed-Effects Model for Mandarin Speakers' Accentedness Ratings in the Image and Voice Match and Mismatch Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	56.69	8.65	[37.60, 75.78]	6.55	.032
Voice-image match	-0.39	4.55	[-9.33, 8.54]	-0.09	.931
Voice-image mismatch	-4.81	4.55	[-13.75, 4.12]	-1.06	.293
Location	7.12	2.37	[2.48, 11.77]	3.00	.003
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	217.00	14.73	Log-likelihood	-1625.20	
Speaker gender (intercept)	111.50	10.56	AIC	3264.40	
			BIC	3291.60	
			Marginal R^2	0.03	
			Conditional R^2	0.56	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels. AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table D10

Summary of Mixed-Effects Model for English Speakers' Accentedness Ratings in the Image and Voice Match and Mismatch Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	5.15	2.13	[1.15, 9.15]	2.42	.045
Voice-image match	0.92	1.83	[-2.66, 4.50]	0.50	.617
Voice-image mismatch	0.40	1.83	[-3.18, 3.98]	0.22	.828
Location	-0.36	0.93	[-2.19, 1.47]	-0.38	.702
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	39.82	6.31	Log-likelihood	-1285.90	
Speaker gender (intercept)	2.89	1.70	AIC	2585.80	
			BIC	2613.00	
			Marginal R^2	0.01	
			Conditional R^2	0.55	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels. AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table D11

Summary of Mixed-Effects Model for Tagalog Speakers' Accentedness Ratings in the Image and Voice Match and Mismatch Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	39.20	22.54	[-14.40, 92.76]	1.74	.324
Voice-image match	-2.41	3.73	[-9.69, 4.88]	-0.65	.527
Voice-image mismatch	0.05	3.73	[-7.24, 7.34]	0.01	.989
Location	0.60	2.05	[-3.36, 4.59]	0.29	.771
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	96.15	9.81	Log-likelihood	-1577.80	
Speaker gender (intercept)	991.14	31.48	AIC	3169.50	
			BIC	3196.70	
			Marginal R^2	0.01	
			Conditional R^2	0.80	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels. AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table D12

Summary of Mixed-Effects Model for Punjabi Speakers' Accentedness Ratings in the Image and Voice Match and Mismatch Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	83.82	3.28	[77.41, 90.23]	25.53	< .001
Voice-image match	-5.35	3.41	[-12.00, 1.30]	-1.57	.118
Voice-image mismatch	-3.81	3.41	[-10.46, 2.85]	-1.12	.265
Location	1.46	1.68	[-1.82, 4.75]	0.87	.386
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	159.28	12.62	Log-likelihood	-1492.30	
Speaker gender (intercept)	0.00	0.00	AIC	2998.70	
			BIC	3025.90	
			Marginal R^2	0.01	
			Conditional R^2	0.66	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels. AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table D13*Summary of Mixed-Effects Model for Combined Accentedness Ratings*

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	72.92	3.75	[64.11, 81.73]	19.45	.019
Mandarin speakers	-15.84	1.38	[-18.54, -13.15]	-11.52	< .001
English speakers	-68.22	1.38	[-70.91, -65.53]	-49.59	< .001
Punjabi speakers	6.70	1.38	[4.01, 9.39]	4.87	< .001
Tagalog speakers	-35.33	1.38	[-38.02, -32.63]	-25.68	< .001
Location	1.60	1.19	[-0.73, 3.94]	1.35	.179
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	55.03	7.42	Log-likelihood	-7907.00	
Speaker gender (intercept)	25.19	5.02	AIC	15832.00	
			BIC	15882.00	
			Marginal R^2	0.64	
			Conditional R^2	0.70	

Table D14

Estimated Marginal Means of Accentedness Ratings for Speaker Groups and Between-Group Comparisons

Speaker group	EMM (1.07)	Speaker group	EMM _{diff} (1.40)	95% CI
Arabic	73.72	Mandarin	15.84	[11.73, 19.96]
		English	68.22	[64.10, 72.34]
		Punjabi	-6.70	[-10.82, -2.58]
		Tagalog	35.33	[31.21, 39.44]
Mandarin	57.87	English	52.38	[48.26, 56.49]
		Punjabi	-22.54	[-26.66, -18.43]
		Tagalog	19.48	[15.37, 23.60]
English	5.50	Punjabi	-74.92	[-79.04, -70.80]
		Tagalog	-32.89	[-37.01, -28.78]
Punjabi	80.42	Tagalog	42.03	[37.91, 46.14]
Tagalog	38.39			

Appendix E

Full Statistical Output for Mixed-Effects Models in Study 2

Table E1

Summary of Mixed-Effects Model for Arabic Speakers' Friendliness Ratings in the Nonreligious and Religious Image Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	78.09	4.53	[69.49, 86.69]	17.24	< .001
Nonreligious image	-0.22	3.78	[-7.63, 7.19]	-0.06	.954
Religious image	0.63	3.78	[-6.78, 8.03]	0.17	.869
Location	1.97	1.83	[-1.61, 5.55]	1.08	.283
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	212.02	14.56	Log-likelihood	-1522.20	
Speaker gender (intercept)	14.40	3.80	AIC	3058.30	
			BIC	3085.50	
			Marginal R^2	0.01	
			Conditional R^2	0.75	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels.

AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table E2

Summary of Mixed-Effects Model for Mandarin Speakers' Friendliness Ratings in the Nonreligious and Religious Image Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	66.07	4.92	[56.52, 75.63]	13.42	< .001
Nonreligious image	6.23	3.88	[-1.35, 13.80]	1.61	.110
Religious image	3.18	3.88	[-4.39, 10.75]	0.82	.413
Location	0.75	1.99	[-3.14, 4.63]	0.38	.705
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	174.11	13.20	Log-likelihood	-1558.30	
Speaker gender (intercept)	20.73	4.55	AIC	3130.70	
			BIC	3157.90	
			Marginal R^2	0.01	
			Conditional R^2	0.54	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels.

AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table E3

Summary of Mixed-Effects Model for English Speakers' Friendliness Ratings in the Nonreligious and Religious Image Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	74.95	4.06	[67.34, 82.55]	18.47	< .001
Nonreligious image	-1.18	3.86	[-8.74, 6.39]	-0.30	.761
Religious image	4.51	3.86	[-3.06, 12.07]	1.17	.245
Location	2.25	1.90	[-1.46, 5.96]	1.19	.236
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	210.14	14.50	Log-likelihood	-1536.00	
Speaker gender (intercept)	5.18	2.28	AIC	3086.10	
			BIC	3113.30	
			Marginal R^2	0.03	
			Conditional R^2	0.69	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels.

AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table E4

Summary of Mixed-Effects Model for Tagalog Speakers' Friendliness Ratings in the Nonreligious and Religious Image Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	74.77	3.91	[67.48, 82.05]	19.11	< .001
Nonreligious image	4.99	3.47	[-1.83, 11.81]	1.44	.152
Religious image	2.47	3.47	[-4.35, 9.29]	0.71	.478
Location	0.96	1.72	[-2.41, 4.34]	0.56	.578
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	161.92	12.73	Log-likelihood	-1503.80	
Speaker gender (intercept)	8.21	2.87	AIC	3021.60	
			BIC	3048.80	
			Marginal R^2	0.01	
			Conditional R^2	0.65	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels.

AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table E5

Summary of Mixed-Effects Model for Punjabi Speakers' Friendliness Ratings in the Nonreligious and Religious Image Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	75.53	4.36	[67.30, 83.76]	17.31	< .001
Nonreligious image	-1.64	3.71	[-8.91, 5.63]	-0.44	.659
Religious image	1.31	3.71	[-5.96, 8.58]	0.35	.724
Location	1.95	1.85	[-1.67, 5.56]	1.05	.293
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	182.68	13.52	Log-likelihood	-1528.90	
Speaker gender (intercept)	12.56	3.54	AIC	3071.80	
			BIC	3099.00	
			Marginal R^2	0.01	
			Conditional R^2	0.64	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels.

AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table E6

Summary of Mixed-Effects Model for Arabic Speakers' Trustworthiness Ratings in the Nonreligious and Religious Image Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	72.26	6.44	[58.61, 85.91]	11.23	.005
Nonreligious image	1.69	3.96	[-6.06, 9.45]	0.43	.669
Religious image	3.66	3.96	[-4.09, 11.42]	0.92	.356
Location	0.89	1.92	[-2.87, 4.64]	0.46	.644
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	233.85	15.29	Log-likelihood	-1539.90	
Speaker gender (intercept)	53.55	7.32	AIC	3093.90	
			BIC	3121.10	
			Marginal R^2	0.01	
			Conditional R^2	0.78	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels.

AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table E7

Summary of Mixed-Effects Model for Mandarin Speakers' Trustworthiness Ratings in the Nonreligious and Religious Image Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	68.53	5.77	[56.82, 80.24]	11.88	.001
Nonreligious image	3.03	4.05	[-4.90, 10.95]	0.75	.456
Religious image	2.68	4.05	[-5.24, 10.60]	0.66	.509
Location	2.79	1.99	[-1.10, 6.69]	1.40	.161
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	230.99	15.20	Log-likelihood	-1554.90	
Speaker gender (intercept)	35.97	6.00	AIC	3123.90	
			BIC	3151.10	
			Marginal R^2	0.01	
			Conditional R^2	0.71	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels.

AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table E8

Summary of Mixed-Effects Model for English Speakers' Trustworthiness Ratings in the Nonreligious and Religious Image Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	71.27	3.75	[64.10, 78.44]	19.00	< .001
Nonreligious image	1.39	3.74	[-5.96, 8.74]	0.37	.710
Religious image	4.61	3.74	[-2.74, 11.96]	1.23	.220
Location	3.30	1.91	[-0.42, 7.03]	1.73	.084
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	168.29	12.97	Log-likelihood	-1540.90	
Speaker gender (intercept)	2.21	1.49	AIC	3095.90	
			BIC	3123.10	
			Marginal R^2	0.02	
			Conditional R^2	0.55	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels.

AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table E9

Summary of Mixed-Effects Model for Tagalog Speakers' Trustworthiness Ratings in the Nonreligious and Religious Image Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	74.91	4.33	[66.77, 83.06]	17.29	< .001
Nonreligious image	2.18	3.71	[-5.10, 9.46]	0.59	.558
Religious image	0.69	3.71	[-6.58, 7.97]	0.19	.852
Location	3.22	1.76	[-0.22, 6.67]	1.83	.068
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	221.40	14.88	Log-likelihood	-1502.80	
Speaker gender (intercept)	11.71	3.42	AIC	3019.70	
			BIC	3046.90	
			Marginal R^2	0.01	
			Conditional R^2	0.85	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels.

AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table E10

Summary of Mixed-Effects Model for Punjabi Speakers' Trustworthiness Ratings in the Nonreligious and Religious Image Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	75.53	5.53	[64.49, 86.57]	13.66	< .001
Nonreligious image	-3.44	4.06	[-11.40, 4.51]	-0.85	.398
Religious image	1.09	4.06	[-6.86, 9.05]	0.27	.788
Location	-0.10	1.96	[-3.93, 3.73]	-0.05	.959
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	250.79	15.84	Log-likelihood	-1545.60	
Speaker gender (intercept)	30.28	5.50	AIC	3105.20	
			BIC	3132.40	
			Marginal R^2	0.01	
			Conditional R^2	0.79	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels.

AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table E11

Summary of Mixed-Effects Model for Arabic Speakers' Intelligence Ratings in the Nonreligious and Religious Image Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	76.08	6.59	[61.97, 90.18]	11.55	.006
Nonreligious image	-0.24	3.92	[-7.92, 7.45]	-0.06	.952
Religious image	0.68	3.92	[-7.00, 8.36]	0.17	.862
Location	-0.25	1.90	[-3.71, 3.71]	0.00	.999
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	229.62	15.15	Log-likelihood	-1536.10	
Speaker gender (intercept)	58.00	7.62	AIC	3086.20	
			BIC	3113.40	
			Marginal R^2	0.01	
			Conditional R^2	0.78	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels.

AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table E12

Summary of Mixed-Effects Model for Mandarin Speakers' Intelligence Ratings in the Nonreligious and Religious Image Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	69.22	4.76	[60.14, 78.29]	14.55	< .001
Nonreligious image	3.63	3.91	[-4.03, 11.28]	0.93	.356
Religious image	2.76	3.91	[-4.90, 10.42]	0.70	.482
Location	1.82	1.99	[-2.08, 5.72]	0.91	.362
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	182.81	13.52	Log-likelihood	-1558.80	
Speaker gender (intercept)	16.92	4.11	AIC	3131.70	
			BIC	3158.90	
			Marginal R^2	0.01	
			Conditional R^2	0.56	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels.

AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table E13

Summary of Mixed–Effects Model for English Speakers’ Intelligence Ratings in the Nonreligious and Religious Image Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	72.45	4.11	[64.43, 80.47]	17.64	< .001
Nonreligious image	0.16	4.24	[–8.13, 8.45]	0.04	.970
Religious image	4.43	4.24	[–3.86, 12.72]	1.05	.297
Location	3.40	1.99	[–0.50, 7.29]	1.71	.089
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	295.71	17.20	Log-likelihood	–1543.30	
Speaker gender (intercept)	0.00	0.00	AIC	3100.60	
			BIC	3127.80	
			Marginal R^2	0.02	
			Conditional R^2	0.87	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels.

AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table E14

Summary of Mixed–Effects Model for Tagalog Speakers’ Intelligence Ratings in the

Nonreligious and Religious Image Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	76.28	4.58	[67.24, 85.32]	16.64	< .001
Nonreligious image	1.58	3.49	[–5.26, 8.43]	0.45	.651
Religious image	1.26	3.49	[–5.58, 8.11]	0.36	.718
Location	–0.11	1.73	[–3.49, 3.27]	–0.07	.948
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	166.84	12.92	Log-likelihood	–1504.20	
Speaker gender (intercept)	19.36	4.40	AIC	3022.40	
			BIC	3049.70	
			Marginal R^2	0.01	
			Conditional R^2	0.68	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels.

AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table E15

Summary of Mixed-Effects Model for Punjabi Speakers' Intelligence Ratings in the Nonreligious and Religious Image Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	74.05	6.31	[60.80, 87.30]	11.75	.003
Nonreligious image	-2.16	4.01	[-10.02, 5.70]	-0.54	.592
Religious image	0.32	4.01	[-7.54, 8.18]	0.08	.937
Location	-0.71	1.94	[-4.50, 3.08]	-0.37	.715
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	243.23	15.60	Log-likelihood	-1542.50	
Speaker gender (intercept)	49.42	7.03	AIC	3099.00	
			BIC	3126.20	
			Marginal R^2	0.01	
			Conditional R^2	0.79	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels.

AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Appendix F

Social Network Survey

For the next task, please think about the people with whom you interact whose **first language is NOT English**. These may be, for example, people with whom you live, work, or share a community. It might be helpful for you to make a list for yourself on a scrap of paper.

When thinking of these interactions, please include people with whom you currently interact **at least once a week** in person, on the phone, or through video chat.

Please also include those with whom you have spoken at least weekly in the year prior to the pandemic, which will give us a more accurate picture of your typical network.

*How many **non-native speakers of English** do you typically interact with weekly through face-to-face conversation, telephone, or video chat?

Please choose... ▾

*For each **non-native speaker** with whom you typically speak weekly (including in the year prior to the pandemic), please indicate the following in the table:

- Age (If you are unsure, please give approximate age.)
- Native language (mother tongue)
- Hours of interaction per week
- Language(s) used in your interaction
- Role (e.g., colleague, physician, neighbour, friend)
- On a scale of 1-5, how close are you with this person? (1 = not very close; you don't know them well, 5 = very close; a very good friend or family member)

Example: 35 Spanish 4 English, Spanish Teacher 3

	Age	Native language	Hours of interaction per week	Language(s) of interaction	Role (e.g., colleague, neighbour, friend)	Closeness (1-5)
Person 1	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Person 2	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Person 3	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

*Which of the following population groups are represented by the non-native speakers of English with whom you interact?

📌 Check all that apply

- Black
- Not sure
- West Asian (e.g., Iranian, Afghan)
- Latin American
- Southeast Asian (e.g., Filipino, Vietnamese, Cambodian)
- South Asian (e.g., East Indian, Pakistani, Sri Lankan)
- Aboriginal (i.e., First Nations, Métis or Inuit)
- White
- East Asian (e.g., Chinese, Korean, Japanese)
- Arab (e.g., Lebanese, Egyptian, Syrian)

*Which religious backgrounds are represented by the non-native speakers of English with whom you interact?

📌 Check all that apply

- Agnostic
- Roman Catholic
- Christian
- Buddhist
- Atheist
- Jewish
- Hindu
- Latter Day Saints or Mormon
- Muslim
- Sikh
- Spiritual, but not religious
- Not sure

Appendix G

Full Statistical Output for Mixed-Effects Models in Study 3

Table G1

Summary of Mixed-Effects Model for Arabic Speakers' Label Choice in the Nonreligious and Religious Image Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	2.66	0.40	[1.78, 3.53]	6.68	.029
Nonreligious image	0.16	0.22	[-0.28, 0.59]	0.71	.481
Religious image	0.39	0.22	[-0.04, 0.83]	1.78	.077
Location	0.09	0.13	[-0.17, 0.35]	0.67	.502
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	0.28	0.53	Log-likelihood	-593.41	
Speaker gender (intercept)	0.23	0.48	AIC	1200.80	
			BIC	1228.00	
			Marginal R^2	0.01	
			Conditional R^2	0.14	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels.

AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table G2

Summary of Mixed-Effects Model for Mandarin Speakers' Label Choice in the Nonreligious and Religious Image Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	3.87	0.33	[3.25, 4.49]	11.60	< .001
Nonreligious image	-0.38	0.30	[-0.97, 0.22]	-1.24	.216
Religious image	-0.33	0.30	[-0.92, 0.26]	-1.10	.274
Location	-0.15	0.15	[-0.45, 0.15]	-0.97	.333
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	1.09	1.04	Log-likelihood	-635.82	
Speaker gender (intercept)	0.05	0.23	AIC	1285.70	
			BIC	1312.80	
			Marginal R^2	0.01	
			Conditional R^2	0.55	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels.

AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table G3

Summary of Mixed-Effects Model for English Speakers' Label Choice in the Nonreligious and Religious Image Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	5.61	0.16	[5.32, 5.91]	35.15	< .001
Nonreligious image	0.01	0.14	[-0.26, 0.28]	0.09	.928
Religious image	0.03	0.14	[-0.25, 0.30]	0.18	.857
Location	0.08	0.08	[-0.08, 0.24]	0.94	.349
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	0.00	0.00	Log-likelihood	-424.11	
Speaker gender (intercept)	0.02	0.13	AIC	862.22	
			BIC	889.43	
			Marginal R^2	0.01	
			Conditional R^2	0.03	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels.

AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table G4

Summary of Mixed-Effects Model for Tagalog Speakers' Label Choice in the Nonreligious and Religious Image Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	4.48	0.97	[2.18, 6.79]	4.62	.121
Nonreligious image	-0.14	0.22	[-0.58, 0.29]	-0.65	.515
Religious image	-0.21	0.22	[-0.64, 0.23]	-0.94	.350
Location	0.13	0.13	[-0.12, 0.39]	1.01	.311
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	0.04	0.19	Log-likelihood	-593.45	
Speaker gender (intercept)	1.80	1.34	AIC	1200.90	
			BIC	1228.10	
			Marginal R^2	0.01	
			Conditional R^2	0.54	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels.

AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

Table G5

Summary of Mixed-Effects Model for Punjabi Speakers' Label Choice in the Nonreligious and Religious Image Conditions

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	2.48	0.21	[2.08, 2.89]	12.00	< .001
Nonreligious image	0.26	0.22	[-0.16, 0.69]	1.21	.232
Religious image	0.18	0.22	[-0.25, 0.61]	0.83	.408
Location	-0.01	0.12	[-0.25, 0.23]	-0.10	.922
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	0.22	0.47	Log-likelihood	-561.30	
Speaker gender (intercept)	0.00	0.00	AIC	1136.60	
			BIC	1163.80	
			Marginal R^2	0.01	
			Conditional R^2	0.17	

Note. The baseline (no image) condition and the Calgary location were set as a reference levels.

AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.