Employability in a Second Language: Recruiting Anglophone Legal *Stagiaires* in Montreal

Madeleine Hyde

A thesis

In

The Department of

Education

Presented in Partial Fulfillment of the Requirements

For the degree of Master of Arts (Applied Linguistics) at

Concordia University

Montreal, Quebec, Canada

August 2015

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Madeleine Hyde

Entitled: Employability in the Second Language: Recruiting Anglophone Legal *Stagiaires*
in Montreal

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

Master of Arts (Applied Linguistics)

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
Sandra Chang-Kredl

_____ Examiner
Pavel Trofimovich

_____ Examiner
Elizabeth Gatbonton

_____ Supervisor
Sarita Kennedy

Approved by: _____

        Graduate Program Director

        _____

        Dean of the Faculty of Arts and Science

Date: _____ 2015

Abstract

Employability in a Second Language: Recruiting Anglophone Legal *Stagiaires* in Montreal

Madeleine Hyde

This study addresses if and how non-native accentedness can impact evaluational judgements of on-paper qualifications in an employment process. This study examines the judgements recruiters make when evaluating job candidates with different types of native and non-native accents. Anglophones who attend university in the officially French-speaking city of Montreal may wish to stay and work after their studies. McGill University's Faculty of Law hosts a bilingual program which is unique in enabling students to work as lawyers in any province of Canada, including Quebec where the language of practice is predominantly French. In this study, highly specialized human resources specialists (Montreal legal recruiters) are asked to evaluate McGill Law students (putatively Anglophone and Francophone) for a particular recruiting process (*Course aux stages*, Montreal's annual "articling" recruitment). The materials in this study simulate, in a controlled environment, a first encounter a recruiter may have with a *stagiaire* candidate. The recruiters are presented with candidates' curriculum vitae (CVs) in conjunction with French-language voice samples from English-accented and French-accented speakers. The voice samples and CVs are counterbalanced, and both quantitative and qualitative evaluational judgements are elicited from the legal recruiter participants. Collection of the data, implications of the results, sharing of qualitative judgements and suggestions for further research are discussed.

Acknowledgements

First and foremost, I would like to thank my supervisor Sarita Kennedy and my committee members Elizabeth Gatbonton and Pavel Trofimovich. Their support and guidance has been integral throughout my thesis project.

I would like to extend much gratitude to the many individuals at the Faculty of Law at McGill University who have supported my academic work. Thank you to my colleagues who helped me pilot my study materials, thank you to Daniel Jutras for supporting the academic endeavors of me as an administrative staff member, thank you to Maryse Chouinard for her help and support, and thank you to the students of the Faculty. The students inspired me to push myself and do a Masters, and then inspired the research questions of this study. This study is for them.

I would like to acknowledge the following individuals: Julie Doré, Andre-Philippe Mallette, Caroline Martel, Mélissa Morand, Nanci Ship, Chima Ubani, and seven other people who have chosen to remain anonymous. I have deep gratitude for their time and valuable expertise.

Thank you to my family and to my dear friends and champions Annette Angell, Julie Beauchamp and Karissa Clayberg. Mostly, thank you to my partner Giovanniantonio Natale who has been my academic inspiration and support system for many years. Ti ringrazio amore.

Table of Contents

List of Tables

Chapter 1

Literature Review

       Languages and cultures have been intersecting at a growing rate over the past century. As the workers of the world migrate in search of new homes and opportunities, employability in a second language is an increasingly relevant area of sociolinguistic research. Employees trying to obtain and retain work in a foreign country (e.g., Canning, 2009; Carlson & McHenry, 2006) or amongst employers who do not use a particular minority dialect (e.g., Atkins, 1993; De La Zerda & Hopper, 1979) can experience linguistic obstacles during the employment process. Discriminatory practices in the hiring process, whether because of gender and parental status, i.e. being a mother, father or childless, (Correll, Benard, & Paik, 2007), visual racial cues (Segrest Purkiss, Perrewé, Gillespie, Mayes, & Ferris, 2006), written racial identifiers, such as typical "Black" and "White" names, (Bertrand & Mullainathan, 2003; McGinnity & Lunn, 2011) or accent (Major, Fitzmaurice, Bunta, & Balasubramanian, 2005; Rakić, Steffens, & Mummendey, 2011; Scassa, 1994) are, unfortunately, prevalent worldwide. Before delving into the specificities of employability and language research, terminology will be defined and a general overview of language attitudes research and its methodologies will be discussed.

       Terminology. In research on language attitudes, the notions of *accent* and *native speaker* are typically defined. For example, the term *accent* can be further unpacked to include the notions of intelligibility and comprehensibility (Munro & Derwing, 1995) and is sometimes confused with other terms, such as pronunciation. Lippi-Green (1997) defines accent as "loose bundles of prosodic and segmental features distributed over geographic and/or social space" (p. 42). In this chapter, accents to refer to the prosodic and segmental features that are distributed in such a way that a listener may identify a speaker of as having a particular language background.

Regarding the term *native speaker,* for the purposes of this study, a native speaker is defined as in Cook (1999): "A person is a native speaker of the language learnt first." (p. 186). Another important characteristic of a native speaker for the current study is someone who has maintained a high level of proficiency in the language into adulthood.

Language attitudes – matched guise technique. The beginning of modern language attitudes research is considered to be in the 1960s, starting with Lambert, Hodgson, Gardner, and Fillenbaum (1960). In their seminal study, Lambert et al. were among the first to employ the matched guise technique (MGT). The MGT is a technique where the *same* person who has complete fluency in two languages or accents produces speech in both languages or accents. The MGT uses the same speaker to control as many extraneous variables in different speech samples as possible. This is done in a way that is not perceptible to the participants, by using distractor voices and interspersing the same person's voices under different guises. In the context of the city of Montreal, Lambert et al. used fully bilingual French-English speakers to elicit evaluation reactions from both native English and French speakers, in order to investigate listener attitudes towards each language. In their study, they found that both English and French participants rated the English voices more favorably, even though, as per the MGT, the French and English voices were actually spoken by the same person.

In the next decade, Seligman, Tucker, and Lambert (1972) furthered language attitudes research by using variety of stimulus materials (photographs, speech samples, written compositions and drawings) to gather subjective impressions from student-teachers about theoretical pupils. In using a variety of materials, they were able to test not just the speech, but also the interaction of the speech with the other attributes the student may have. They found that

voice and photograph cues significantly impacted the student-teachers' attitudes and impressions of the students' intelligence.

In the 1980s, more studies were conducted using the MGT to investigate language attitudes, that used increasingly nuanced methodologies (Ball, 1983; Seggie, Fulmizi, & Stewart, 1982). Ball's work was significant as he raised an important methodological consideration: when using the traditional MGT, listener boredom may occur when listening to repetitive samples, which draws people to pay attention to the non-content (i.e., the speech). They used this to their advantage to elicit their reactions to the *voice* in the samples. Seggie et al.'s work makes another contribution, which is that a perceived "educated" speaker may be perceived as less suitable for a high-status job than someone who has a perceived less "educated" accent. With the contribution of Seggie et al.'s research, subsequent researchers (Carlson & McHenry, 2006; Hosoda, Nguyen, & Stone-Romero, 2012) would be inspired to further investigate attitudes towards accentedness using different levels of "status" in their materials, job-related or otherwise.

To conclude, the use of the MGT over the past half-century has consistently shown that listeners have different reactions to the same speech content spoken in different accents, even when the speaker is, in fact, the same. This signifies that even when voices are "matched," the style, accent or dialect of speech will colour the listeners' perceptions.

An alternative to the matched guise technique – multiple speakers. Other researchers have opted to conduct similar studies with techniques other than the matched-guise. In Kalin and Rayko (1978), the researchers were among the first to use various foreign-accented speech samples, spoken by different voices, to elicit spontaneous attitudinal reactions in an employment context. Later, Atkins (1993) conducted research on Appalachian and Black English that was not only un-matched, but written, not spoken. Rather than using the same people for her samples and

matching them, she used representative written samples from both Appalachian and Black English speech groups. These samples were representative because they used speech conventions of both Appalachian and Black English. For example, in Appalachian English one might say: "I knowed it" instead of "I knew it" (p. 112) and in Black English one might say: "aks" instead of "ask" (p. 113). In Atkins' work, however, she did not elicit evaluational reactions; rather, she asked her participants direct questions about the way the candidate was using language and if the written dialect was having an effect on what the participant thought. For example, Atkins asked participants to agree or disagree with statements such as, "even if this person fulfilled other job qualifications and requirements, I definitely would not hire him/her because of this dialectal characteristic," (p.110). However, written text has conventions which are substantially different from those of oral speech, so the evaluation of written-down oral speech is not wholly authentic. While this research method seems somewhat inelegant and inauthentic, Atkins' results are consistent with other studies that elicit evaluational reactions; her participants rated both minority "speech" samples negatively when being rated for various employability characteristics.

More convincingly, Markley (2000) enlisted the use of various USA regional speakers for her study, rather than using a matching technique. She pretested for representativeness of each voice and controlled the samples by having her speakers recite the same passage. Markley opted for male voices who had no distinct features in their voice to indicate age. Her results, as with other research that employs or does not employ the MGT, consistently show that different styles of speech result in different evaluations.

Language attitude research – stereotyping and discrimination. Research from the past half century show that listeners' attitudes have a significant effect on perceptions. Research is often conducted through indirect means (such as the MGT), because attitudes cannot be

measured directly. Indirect measures may be used to investigate language stereotyping and other discriminate behaviors that people may not be aware they have. For example, Rubin and Smith (1990) found that undergraduate students perceived nonnative English speaking teaching assistants to be poor teachers when they perceived a high level of accentedness. More troubling still, studies have found that teachers with perceived non-native accents are more likely to face discrimination, especially when considered alongside native speaker teachers (Clark & Paran, 2007; Medgyes, 1992). Language discrimination and stereotyping can have a drastic impact not only on being a teacher and general employability (Clark & Paran, 2007; Hosoda et al., 2012; Parton, Siltanen, Hosman, & Langenderfer, 2002) but also on workplace segregation (Hellerstein & Neumark, 2008) and even on the ability for someone with a detectable accent to find a place to live (Purnell, Idsardi, & Baugh, 1999). In Purnell et al.'s study, they found, by conducting a phone survey, that Hispanic speakers were discriminated against when trying to phone landlords to enquire about being a tenant. These studies tested behaviors indirectly, to gather information about attitudes that people have that they may not admit to, if asked outright. As the research shows, language stereotyping and discrimination has far-reaching and varied impacts on the lives of people with perceived accentedness.

Listener preferences and perceptions in regards to accent, pronunciation and comprehensibility have been investigated in both the Canadian-specific context (Derwing & Munro, 2009; Munro & Derwing, 1995) and also in the United States (Lindemann, 2002, 2010). Lindemann (2002) found that those with predetermined negative attitudes towards Koreans had poorer comprehension of Korean-accented speakers. This is significant to keep in mind when reviewing research conducted about language attitudes and listener preferences, as Lindemann

reminds researchers that communication is reciprocal and that the *listener* must be open to different accents and styles of speech in order for communication to be a success.

Language and employability – onus on the interviewee. While some researchers (Gluszek & Dovidio, 2010; Lindemann, 2002, 2010) remind us that communication is reciprocal, in an employment-search situation in today's economy, the onus of making a positive impression typically falls on the prospective employee, often during the interview process. Biases, whether known or not to the interviewer, interfere with a candidate's chances of obtaining employment (Davison & Burke, 2000; Pingitore, Dugoni, Tindale, & Spring, 1994). For candidates who do not speak the dominant language or dialect as a native speaker, employment prospects can be affected by the interviewers' evaluational judgements that may arise from perceiving a particular language background (or, accent) in candidates' speech.

Employability research: CVs and voice samples. No prior research has paired various accented voice samples with CVs to elicit evaluational reactions, as in the current study. The closest such research is that of Segrest Purkiss et al. (2006) who used a video recording of an interview together with "participant demographic information" (prior work experience, GPA, etc.), which is information similar to what can be found on a standard CV. In viewing this participant demographic information, the raters were able to make more informed employability judgements when watching the video recording of the interview. Many studies have used voice samples as the sole materials, and some more recent studies have used CVs alone to test for potential discriminatory behavior (Bertrand & Mullainathan, 2003; Correll et al., 2007; McGinnity & Lunn, 2011). However, studies which combine voice samples and some sort of information about the candidate give a more complete portrait of the candidacy in regards to accent and employability. As well, when researchers want to test whether participants' accents

affect their employability, a logical and more authentic choice of materials would be samples of their voices and samples of their employment qualifications (CVs). Therefore, in designing the current study, the methodology and implications of both studies that use voice samples and of studies that use CVs as the materials were considered, in order to create a single study combining both kinds of materials.

Employability research: audit studies. An audit study "combines experimental design with real-life settings… Distinct from more laboratory studies, audit study participants are the people who make important decisions about actual applicants," (Correll et al., 2007, p. 1327). Audit studies are useful because they can canvas a wide variety of authentic rater participants, which allows researchers to gather authentic reactions and judgements from genuine employers. Bertrand and Mullainathan (2003) conducted an extensive audit study, where they sent 5000 resumes for over 1300 real job postings in two different American cities to test to see if the name on the CV had an effect on the call-back rate. They manipulated the names on the CVs to be typically African American or White, and mixed and matched the content of the CVs with the names. The CVs were created using real CVs as inspiration, and they were manoeuvred into being "high" (strong) or "low" (weak) in content. Bertrand and Mullainathan found that the names on the CVs had a statistically significant effect, as the CVs with the White name condition were fifty percent more likely to receive a call-back than a CV with an African American name. Later, McGinnity and Lunn (2011) also conducted an audit study in Ireland by investigating preference of Irish-sounding names over Asian, German or African-sounding names on CVs which were sent to actual employers. Their results aligned with those of Bertrand and Mullainathan's, as their Irish-sounding CVs were also twice as likely to receive callbacks as equal CVs with non-Irish sounding names. Correll et al. (2007) conducted both a laboratory and

an audit study similar to those of Bertrand and Mullainthan and McKinnity and Lunn. They manipulated CVs to test for discrimination in the context of race (White vs African American), gender, and parental status, with a particular focus on parental status. Correll et al. did in fact find discriminatory behavior towards women who presented themselves as parents.

Authenticity – study design and targeted materials. Correll et al.'s work is important for employability research, as it mixed both laboratory and audit research when testing for the same variables, rather than choosing just one approach to conducting research. Their work is more convincing than those who chose one approach (laboratory or audit) because in using both approaches, they were able to balance each approach's shortcomings against each other. In other words, the inauthentic environment of laboratory studies was balanced by the more authentic environment of audit studies. Conversely, the lack of a controlled environment in audit studies is balanced by the controlled environment in laboratory studies. Correll et al. also fabricated a convincing cover-story when conducting the laboratory research with their undergraduate participants, providing a compelling reason for completing the rating task. The students were told that a start-up communications company was looking for the insight of young people as heavy consumers of technology. They were also told that their input would affect future hiring decisions, in order to give the participants a sense of purpose and responsibility, thus, possibly gathering more thoughtful data.

As well, Correll et al., Bertrand and Mullainthan, and McKinninity and Lund all made great efforts to ensure that their CVs were balanced and as authentic as possible. Authentic-like CVs mean that, especially when conducting an audit study, those who receive the documents are convinced of their authenticity. Researchers would therefore hope that the employability reactions from the employers are also authentic. Since CVs are, by definition, rich in content,

careful construction and counterbalancing of the CVs is of the utmost importance to ensure that extraneous factors on CVs are not affecting rater perceptions and hence, the results (Correll et al., 2007).

Many researchers seem to discount the importance of tailored employment materials. In the case of Correll et al. (2007), the researchers included cover letters along with fake applications sent to hundreds of jobs. Although they didn't report many specifics of the cover letters, from what they did report, it seems unlikely that they spent hours tailoring each cover letter for each specific job posting. In today's economy, across job fields, it is recommended to spend time writing specific cover letters for the job and place of work. Recruiters, as trained professionals, can see when application documents are generic and not tailored to the specific position available. This means that researchers attempting to discover the implication of bias in an authentic recruiting context won't be able to convincingly apply these results to job-seekers who take the time to create tailored cover letters. In the current study, the use of tailored materials representative of a specific pool of job-seekers means that the raters are presented with materials which are authentic and relevant for their typical employment evaluation scenario, thus encouraging judgements which are representative of raters' real-life evaluations.

Overall, when conducting employability research, researchers should endeavor to recreate as realistic a simulation as possible by using convincing cover stories and authentic employment materials. Researchers should also try to use people who make hiring decisions as participants.

The current study. Making use of specialized raters in an authentic assessment context, my goal in this study is to determine how recognizable L2 Anglophone accents in French might affect employment decisions. The participants in this study are legal recruiters from the Montreal

area who make hiring recommendations and decisions regarding an annual *Course aux stages*, *stagiaire* ("articling" student) recruitment process. Since the dominant language of legal employment in Montreal is French, this study investigates whether and how candidates' non-native French accents may affect the perceptions of possible employers in terms of candidates' suitability. This study creates realistic candidate profiles by combining a range of curriculum vitae (CVs) with speech samples from female candidates with native or non-native French accents, answering a common interview question in French. The ways in which raters perceive these combinations of CVs and speech samples should provide insight into whether and how candidates' accents affects perceptions of their employability.

The following chapter is a manuscript-length report on the research study. The manuscript begins with a literature review about L2 users and employment, a problem statement, followed by description of the methodology and results, ending with discussion of the findings and conclusions. A concluding chapter following the manuscript contains general conclusions and discussions, suggestions for further research, the references and appendices.

Employability in a Second Language: Recruiting Anglophone Legal *Stagiaires* in Montreal

Literature Review

As the world becomes more globalized, employability in a second language is an increasingly relevant area of sociolinguistic research. In the current study, I investigate whether job-seekers' non-native accents play a role in hiring decisions. Located in the French-English bilingual city of Montreal, Canada, this study simulates the first phase of a particular hiring process. In this study, the job-seekers – Anglophone and Francophone students – identified as having a specific educational background – McGill University, Faculty of Law – are evaluated for specific positions – an entry-level legal position (a *stage*). This study uses an experimental design which mixes different accented voice samples and different levels of qualifications (curriculum vitae).

Employees trying to obtain and retain work in a foreign country (e.g., Canning 2009; Carlson & McHenry, 2006) or amongst employers who do not use a particular minority dialect (e.g., De La Zerda & Hopper, 1979; Ball, 1983), can experience linguistic obstacles during the interview process. Discriminatory practices in the hiring process such as: applicants' parental status, (Correll et al., 2007); visual racial cues (Segrest Purkiss et al., 2006); racial identifiers, such as typical "Black" and "White" names, (e.g., Darity & Mason, 1998; Bertrand & Mullainathan, 2003); or accent (e.g., Atkin, 1993; Rakic et al., 2011) are, unfortunately, prevalent worldwide. Some researchers (Gluszek & Dovidio, 2010; Lindemann, 2002, 2010) remind us that communication is reciprocal and that the listener must be open to different accents and styles of speech in order for communication to be a success. However, in the current economy, the onus of making a positive impression typically falls on the prospective employee,

often during the interview process. Biases, whether known or not to the interviewer, can interfere with a candidate's chances of obtaining employment (e.g., Davidson & Burke, 2000; Pingitore et al., 1994). For candidates who do not speak the dominant language or dialect, employment prospects can be affected by interviewers' known or unknown biases and subsequent judgements upon hearing a particular accent.

Terminology. In research on employability in a second language or in a minority-dialect, the terms *accent* and *native speaker* are typically defined before moving forward. For example, the term *accent* can be further unpacked to include the notions of intelligibility and comprehensibility (Munro & Derwing, 1995) and is sometimes confused with other terms, such as pronunciation. Lippi-Green (1997) defines accent as "loose bundles of prosodic and segmental features distributed over geographic and/or social space" (p. 42). In this study, *accent* refers to the prosodic and segmental features that are distributed in such a way that a listener may easily identify a speaker of French as being from either an Anglophone or Quebecois Francophone background. In Quebec, Canada, a particular variety of French is spoken, which is called *Quebec French* or *Quebecois.* (For more information about attitudes towards Quebec French, see Salien, 1998). Regarding the term *native speaker,* for the purposes of this study, a native speaker is defined as in Cook, (1999): "A person is a native speaker of the language learnt first." (p. 186). Another important characteristic of a native speaker for the current study is someone who has maintained a high level of proficiency in the language into adulthood.

Language attitudes. Starting in the 1960s with the seminal study by Lambert et al. (1960), researchers have been investigating reactions and attitudes towards second language speakers and accents. Lambert et al. were among the first to employ the matched guise technique, (which will be explained further in the next sections) by using bilingual French-

English speakers to elicit evaluation reactions from both native English and French speakers, in order to investigate attitudes towards each language in the context of the city of Montreal, Canada. In their study, they found that both English and French participants rated the English voices more favorably, even though, as per the MGT, the French and English voices were the same person speaking. In the next decade, Seligman et al. (1972) used a variety of stimulus materials (photographs, speech samples, written compositions and drawings) to gather subjective impressions from student-teachers about theoretical pupils, finding that voice and photograph cues significantly impacted listeners' attitudes and impressions of speakers' intelligence. Later, in the 1990s, (Rubin & Smith, 1990) found that undergraduate students perceived nonnative English speaking teaching assistants to be poorer teachers when they perceived a high level of accentedness. Finally, (Lindemann, 2002) found in her study that those with a predetermined negative attitude towards Koreans had poorer comprehension of Korean-accented speakers, which demonstrated that the *listener* aspect of communication can be just as important as the *speaker*. Thus, the research from the past half century shows that language attitudes has a great effect on many different types of perceptions and judgements.

Stereotypes, discrimination and language. Along with language attitudes comes language stereotyping and opportunities for potential discrimination. Language discrimination and stereotyping can have a drastic impact not only on employability (Hosoda et al., 2012; Parton et al., 2002) and workplace segregation (Hellerstein & Neumark, 2008) but also on the ability for someone with a detectable minority accent to find a place to live (Purnell et al., 1999). Along with investigating the speaker, *listener* preferences and perceptions in regards to accent, pronunciation and comprehensibility have been investigated in both the Canadian-specific context (Derwing & Munro, 2009; Munro & Derwing, 1995) and in the United States

(Lindemann, 2002, 2010). Several techniques have been used to investigate discriminatory language stereotyping behaviors, one of which is the matched guise.

Methodologies used – matched guise technique. Much accent-based research of discriminatory practices and judgements has used the "matched-guise technique" (MGT), which is a technique developed by Lambert et al. (1960) that uses the same speaker to control as many extraneous variables in different speech samples as possible. This is done in such a way that is unobvious to the participants, by using distractor voices and interspersing the same person's voices under different guises. Use of the MGT over the past half-century has consistently shown that different evaluational reactions can be made on the same speech content spoken in different accents, even when the speaker is, in fact, the same (e.g., Hosoda et al., 2012; Lambert et al., 1960). This signifies that even when voices are "matched," the style, accent or dialect of speech colours the evaluator's perception.

Gaies and Beebe (1991) and Laur (2014), however, question the validity of the technique. Laur (2014) distrusts whether, given the research, the MGT achieves the validity many researchers claim:

> The very nature of the matched guise method – twinning language pairs – contains the possibility of bias through imbalance. A linguistic pair provides information on how the elements of each language compare with the other, but not… on other characteristics that could have a direct effect on results (p. 21).

As well, the MGT, by design, must use non-random and recorded samples, which are not spontaneous or authentic. Laur claims that even if there is a significant relationship found when using the MGT, we cannot be certain of any relationship beyond the one individual who speaks with both accents, as the relationship will only be between the speaker and himself. Whereas,

with two or more speakers with different accents, the study can evaluate if raters react to two speakers differently. Therefore, while the MGT is one technique to test for potential judgements and biases, it is not without its imperfections and imbalances.

Methodologies used – multiple speakers. Other researchers (e.g., Atkins, 1993) have opted to conduct similar research with techniques other than the MGT. In a study conducted by Kalin and Rayko (1978), the researchers were among the first to use various foreign-accented speech samples, spoken by different voices, to elicit spontaneous attitudinal reactions. As well, Markley (2000), enlisted the use of various USA regional speakers for her study instead of using a matching technique. Their results, as with similar research that employs or does not employ the MGT, show that different styles of speech, regardless of the content, result in different evaluational judgements.

CVs and voice samples. No prior research has paired various accented voice samples with CVs to elicit evaluational reactions, as in the current study. The closest such research is that of Segrest Purkiss et al. (2006) who used viewing a video recording of an interview together with "participant demographic information" (prior work experience, GPA, etc.), which is information similar to what can be found on a standard CV. In viewing this participant demographic information, the raters were able to make more informed employability judgements when watching the video recording of the interview.

While some of the aforementioned studies used voice samples as the sole materials, some more recent studies used CVs alone to test for potential discriminatory behavior (Bertrand & Mullainathan, 2003; Correll et al., 2007; McGinnity & Lunn, 2011). Such studies cannot test for accent, so when investigating if someone's accent is impacting their chances of employability, the logical choice of materials would be a sample of their voice, and a sample of their

employment qualifications (i.e., a CV). Therefore, in considering the current study's design, studies which used voice sample and studies which used CVs to explore discriminatory behaviour were reviewed. This was done to identify the beneficial features of both types of materials, which will be further examined following the discussion of participants.

Participant selection. When testing for employability, researchers tended to recruit one of two types of participants: undergraduate students who were instructed to pretend they were recruiters (e.g., Cargile, 2000; Giles et al., 1981) or recruiters themselves (e.g., Atkins, 1993; Ugbah & Evuleocha, 1992). Correll et al. (2007) and Parton et al. (2002) integrated the use of both recruiter and student participants. Laur (2014) believes the use of undergraduate students is merely convenience sampling, as students are often readily willing and available for studies in exchange for course credit or stipends. Results gathered from student participants are not as convincing because students are not typically experienced enough to make representative judgements from the perspective of an employment professional. These data are not as useful, because to survey a populace who are not recruiters is not authentic or applicable to real-life employment decisions. With the exception of Correll et al. (2007), which will be discussed in the next sections, and Kalin and Rayko (1978), the use of students or random raters and their particular demographic is unaddressed or unaccounted for in the design of many studies.

Audit studies. An audit study "combines experimental design with real-life settings… distinct from more laboratory studies, audit study participants are the people who make important decisions about actual applicants," (Correll et al., 2007, p. 1327). Audit studies are valuable because they can canvas a wide variety of authentic participants, which allows researchers to gather authentic reactions and judgements from genuine employers. Bertrand and Mullainathan (2003) conducted an extensive audit study, where they sent 5000 resumes to over

1300 real job postings in two different cities to test to see if the name on the CV had an effect on the call-back rate. They manipulated the names on the CVs to be typically African American or White, and mixed and matched the content of the CVs with the names. Bertrand and Mullainathan found that the names on the CVs did have a statistically significant effect, as the CVs with the White name condition were fifty percent more likely to receive a call-back than a CV with an African American name. Correll et al. (2007) conducted both a laboratory and an audit study similar to that of Bertrand and Mullainthan. They manipulated CVs to test for race, gender and parental status discrimination, and they did in fact find discriminatory behavior towards women who presented themselves as parents in both the laboratory and audit parts of their study.

Authenticity – study design and targeted materials. Correll et al.'s work is important for employability research, as it mixed both laboratory and audit research when testing for the same variables, rather than choosing just one approach to conducting research. Their work is more convincing than those who chose one approach because the inauthentic environment of the laboratory research was balanced by the more authentic environment of the audit research, while the lack of a controlled environment in audit research was balanced by the controlled environment in the laboratory research. Correll et al. also fabricated a convincing cover-story when conducting the laboratory research with their undergraduate students participants, providing a persuasive reason for completing the study, and thus eliciting possibly more thoughtful reactions.

Many researchers (e.g., McGinnity & Lunn, 2011; Bertrand & Mullainathan, 2003) made great efforts to ensure that their CVs were balanced and as authentic as possible. Authentic CVs means, especially in an audit study, that those who receive the documents are convinced of their

authenticity; researchers would therefore hope that the employability reactions from the employers are also authentic. Since CVs are, by definition, rich in content, careful construction and counterbalancing of the CVs is of the utmost importance to ensure that extraneous factors on CVs are not affecting rater perceptions and hence, the results (Correll et al., 2007).

Overall, when conducting employability research, researchers should endeavor to recreate as realistic a simulation as possible by using convincing cover stories and authentic employment materials. Researchers should also try to use people who make hiring decisions as participants.

Problem Statement

Job-seekers who are applying for employment opportunities in a second language should be aware of whether their accents may affect their chances of gaining employment. If their accents do affect their chances, they should understand the individual characteristics or experiences which might mitigate the effect of a recognizable nonstandard accent. In job interview settings, personal presentation, including speech, is crucial. Employers often provide little time for a first interview, and first impressions can take mere seconds to form (Dougherty, Turban, & Callender, 1994). A second language accent may be perceived within moments, and can be interpreted many different ways. A review of past and contemporary research shows that while employability and language usage have been researched before in many different ways, there is no published study that considers employability of Anglophones when trying to gain employment in another language. This is unsurprising, as a great many L2 job-seekers worldwide try to find employment in English, which is the lingua franca for many disciplines. However, considering the number of students who come to study at universities in the officially French-language province of Quebec from other, English-speaking, provinces, (not to mention

the United States and other countries), this gap in the research should be filled, as conducting research for this demographic could be useful for students when assessing their employment prospects in Quebec. This data could also be used by Montreal universities in presentations to prospective students who wish to come not only to study, but to stay and make a life in the city.

Students in professional university programs in English-medium universities in Quebec typically receive training which meets professional standards for employment in Quebec or in other provinces; however, this may not include language training for the workplace. For example, students entering the Bachelor of Civil and Common Laws program in the Law Faculty of McGill University must be at least "passively bilingual." For the Faculty, this means being able to attend lectures and read in both languages (English and French), but not necessarily to produce language by speaking or writing. Anglophone law students who seek entry-level "articling" positions (*stages)* in Quebec are sometimes concerned that, during the recruitment process, potential employers may be negatively affected by students' demonstrably non-native use of French. Although there is anecdotal evidence that Anglophone students have been successful in the annual *Course aux stages* recruitment process*,* no existing research has examined whether or how students' non-native use of French affects the decisions of potential employers.

The Current Study

Making use of specialized raters in an authentic assessment context, my goal in this study is to determine how L2 Anglophone accents in French which are perceptible to recruiters might affect employment decisions. The participants in this study are legal recruiters from the Montreal area who make hiring recommendations and decisions regarding an annual *Course aux Stages*, *stagiaire* ("articling" student) recruitment process. Since the dominant language of legal

employment in Montreal is French, this study investigates whether and how candidates' non-native French accents may affect the perceptions of possible employers in terms of candidates' suitability. This study creates realistic candidate profiles by combining different strengths of curriculum vitae (CVs) with speech samples from female candidates with native or non-native French accents, answering a common interview question in French. The ways in which raters perceive these combinations of CVs and speech samples provides insight into whether and how candidates' accents affect perceptions of their employability.

Research Questions and Hypotheses

Research Question 1: Which aspects of a *stagiaire* candidate's profile, as demonstrated in their CVs, are valued by recruiters for employability purposes?

Research Question 2: When considering potential *stagiaire* candidates, how is the recruiter's perceptions of candidates' employability, considering also CVs and GPAs, affected by candidates' L2 accents when speaking French?

Research Question 3: When recruiters use CVs, GPAs and oral responses to judge the weakness or strength of speakers' candidacy, are their evaluations mediated by their judgments of the strength or weakness of candidates' CVs and GPAs?

Hypotheses

Hypothesis 1: Candidates with a strong CV and GPA will be rated favorably regardless of accent.

Hypothesis 2: Candidates with a weak CV and GPA will be rated more negatively – though not overwhelmingly negatively – if candidates have an Anglophone accent.

Hypothesis 3: Candidates with an average CV and GPA will be rated favorably regardless of accent, though less favorably than candidates with strong CVs and GPAs.

Null hypothesis: No significant relationships will be found between raters' evaluations and any specific type of CV, GPA, or accent.

Methodology

Rater Participants

Thirteen rater participants, who were all legal recruiters in the Montreal area from medium-to-large sized private law firms and legal governmental organizations, participated in this study. There were ten female and three male participants. All participants, except one, studied law and were members of a legal Bar association. Of the participants that studied law, two studied at McGill University, and ten did not. Two participants identified having English as a native language, and eleven identified having French as a native language. More information about the participants and their language usage and expertise is found in Table 1.

This pool is small in comparison with studies that use students as raters; however, there are a select number of legal recruiters with the resources to recruit *stagiaires* annually in the *Course aux stages* process. In 2015, there were 24 employers who signed the annual Montreal-wide recruitment agreement, although there are also a small number of employers who follow the recruitment agreement unofficially. Therefore, the number of raters in this study is not only appropriate, but also aligns with other research done in a specific employment context (Derwing & Munro, 2009).

Table 1
Rater participant information

Characteristics	Yes	No
Born in Quebec	12	1
Member of a Bar Association	12	1
Conducted a minimum of 10 interviews in McGill *Course aux stages* process	12	1
Firm/organization part of *Course aux stages* process *or* follows similar timeline to process	13	0
Conduct legal recruitment interviews on a regular basis	13	0
Review CVs on a regular basis	13	0
Work in private practice	12	1
Studied law at McGill	2	11
Head of *stagiaire* recruitment of firm/organization	11	2
Average number of languages spoken (range)	2.3 (2-3)	
Average self-perceived ability in French on scale of 1(weak) to 6 (strong)	5.8	
Average self-perceived ability in English on scale of 1(weak) to 6 (strong)	5.1	
Average need to **speak** French at a high level at workplace on scale of 1(weak) to 6 (strong)	5.4	
Average need to **write** in French at a high level at workplace on scale of 1(weak) to 6 (strong)	5.2	
Average percentage of workday using English	37.5	
Average percentage of workday using French	62.5	

Materials

The materials in this study were created by me, the author. The first contact legal

recruiters have with candidates is often on paper (or, nowadays, electronically) via candidates'

application materials, which typically consist of transcripts, a CV and a cover letter. In addition,

legal recruiters might briefly meet a candidate at an informal event, such as a career fair or a

networking cocktail. Therefore, the two materials in this study serve to simulate a first encounter

with a hypothetical candidate: voice samples and CVs. There is also a third element to the

materials, which is the content of the speech samples and CVs. The accents of the voices, the

content of what is being said, and the content of the CVs have been mixed and counterbalanced

across participants. The voice samples were 75-word responses in French to a standard interview

question ("How do you react to criticism?"), which served to provide meaningful and representative context to hear the voices. The CVs, which included GPAs, provided further context in which to evaluate the candidate. Cover letters were not included, because of the limited time the participants had available for rating and to avoid potential threats to validity because of differences in content or writing style in the letter. Transcripts were not included because of confidentiality concerns. The law GPA and a mention of undergraduate academic accomplishments were, however, included, so that the participants had an indication of academic abilities. As CVs are fairly standardized in entry-level legal recruiting, including this material allowed the participants to make judgements of the candidate based on more than just a short voice sample.

The curriculum vitae. The three CVs were between 1.75-2 pages in length. One CV was "Weak," one "Average," and one "Strong." Appendix A provides a detailed breakdown of what was included in each CV. All CVs were written in English, but included knowledge of French in a "languages" section. Each CV included the students' law GPA. For the actual *Course aux Stages* process, students may choose which language to write their CV in. Based on my experience as Career Development Officer in the McGill Faculty of Law, students of both French and English language backgrounds choose to write their CVs in either French or English, regardless of native language.

The dates on the CVs showed that candidates went directly into their legal studies after finishing another degree and did not take time off to work or pursue other interests. The CVs included entry-level work experience during undergraduate studies. Each CV included service-industry work experience as well as university work experience. The CVs had some international experience, as is common amongst McGill law students. Hobbies and interests were included,

which comprised a balanced proportion of intellectual, individual, and group interests. As part of my job duties as Career Development Officer at McGill University, I have reviewed authentic CVs of hundreds of McGill law students; the CVs in the current study were inspired by this experience, although they do not reflect any one person in particular.

To be societally and linguistically neutral, and per the findings of Bertrand and Mullainathan (2003) and McGinnity and Lunn (2011), the CVs did not include names or addresses. Rather, the CVs were coded by number, for example, "CV 1." Appendix A breaks down, in general order of appearance on the CV, what has been included on each CV.

The Average CV. The Average CV was approximately 1.75 pages in length, which is a normal length of a Canadian law student's legal CV. The GPA was in the top 50th percentile of the class ("McGill Faculty of Law: Grade Distribution Table," 2015) which at McGill Law typically falls between 3.0 and 3.2 on a 4.0 scale. This is generally considered strong enough to warrant an interview in the *Course aux stages* process, if the other aspects of the candidate profile are also compelling.

The Weak CV. The Weak CV contained fewer words and less substance. The length was not meant to be starkly different, rather, subtly different enough to suggest that the candidate had less experience. The undergraduate degree was not at an honours level, suggesting that the student took a less rigorous course load. Most importantly, the law GPA was manipulated into being just below 3.0 (which is a B average at McGill). Many students and some employers view 3.0 as an invisible barrier between an "acceptable" and an "unacceptable" GPA. For certain McGill law programs, such as the highly competitive USA recruitment process and clerkships with a judge for academic credits, 3.0 can sometimes serve as a suggested "cut-off" for

participation. While this CV contained work experience during prior studies and in the summer, the experience was less ample and prestigious than the Average and Strong CVs.

The Strong CV. The Strong CV was slightly longer than the Average CV. The candidate had, in addition to an honours undergraduate degree, a graduate degree. The candidate had won a McGill Law entrance scholarship, which signifies excellence in academia. In conjunction with this, the candidate's GPA was manipulated into being approximately in the top 15% of the class (roughly a 3.30 GPA). This GPA is considered strong by many employers, including those who hire during the *Course aux stages* process, who are aware of the rigours of the McGill Law grading system. It is not, however, the top-most percentile of the class, which could facilitate a first interview in and of itself, despite any other factors in a candidate's profile. The student's previous summer work experience in the summer after their first year of law school included work in a legal environment, as it is very difficult for students to gain directly relevant legal employment after only one year of study.

The Voice Samples

Response content. The voice samples contained content, spoken in French, in response to the standard interview question, "How do you react to criticism?" (*Comment réagissez-vous à la critique?*). The response content was a collaborative creation of the author and a native French-speaking professional who has extensive experience in interviewing law students. The question and response content were chosen based on the mock interviews held by the Career Development Office at McGill University's Faculty of Law.

In order to mitigate threats to validity due to the content of the speech sample and not the accent of the person speaking, there were three different versions of the answer, all of which were 75 words in length (or, between 24-29 seconds of speech). Based on my experience

working in the Faculty of Law, McGill law students tend to answer this banal question in a similar fashion, but using different words and expressions. This was ideal for the content of the answer having as neutral an effect as possible on the judgements made, as the goal was to test for accent, not content of response. The three versions of the answer were piloted, first in written and then in spoken form, to ensure that the content of the responses were rated as similar as possible on various scales, such as: how well the response answers the question, how appropriate the response content is, the overall ranking of the response, etc. In piloting the written samples, native speakers of French with knowledge of McGill law students in their roles as students or staff members were asked several questions about the response content and how well the response answered the question, along with questions about general language use and style. The pilot test participants were also asked to rate the strength of the answer and provide general commentary. Out of the four response contents created, one was eliminated based on the feedback, and the other three were rated similarly enough and slightly adjusted based on feedback, for a total of three final response contents.

The voices. There were three voices: one native French speaker (Quebecois French) and two North American Anglophones who had self-assessed high levels of proficiency in French. All of the speakers were female, 26-27 years old, and read from the aforementioned carefully constructed scripts. The French speaker ("FF" – Quebecois French accent) was a native Quebecer who grew up near Montreal and had lived in Montreal for most of her adult life. One of the Anglophone speakers ("FE" – Anglophone accent 1) grew up in Western Canada and went through the French Immersion stream for part of her schooling years. While she lived in Montreal at the time of recording, she didn't use French extensively or on a daily basis and has had inconsistent oral French practice in her adult years. The other Anglophone ("FM" –

26

Anglophone accent 2) grew up in the Eastern part of the United States, with limited French language schooling during childhood. In adulthood, she studied French as a university major in both undergraduate and graduate programs in Montreal, and at the time of recording spoke French on a daily basis with her partner, who was Quebecois, with friends and at work.

Pilot testing of the voices speaking the response content was conducted with both native French and native English speakers, all of whom can understand spoken French. In the pilot testing, the participants were asked: what native language they thought the speaker had, if the speaker's accent interfered with comprehension, how accented the speaker sounded, and, after each question, how confident they were of their own judgements. In the piloting, voice FF was consistently rated as native Francophone, FE was consistently rated as Anglophone, but there was some lack of confidence as to FM's language background being Anglophone, particularly with fellow Anglophones. FM's accent was judged to be less Anglophone accented as FE. Therefore, the two voices with different degrees of non-native accentedness (FE and FM) were included in order to explore whether a degree of accentedness made a difference for recruiters' evaluations.

Procedure and Design

Each rater participant was given a personal profile questionnaire that was anonymously linked to their rating responses. The questionnaire asked the participant what sort of firm or organization they worked for, what their native language is, how much they interacted in both French and English at their place of work, and more. The participant then made evaluational judgements of job candidates based on a provided CV, using binary and 6-point Likert-type scales (see description of procedure in Appendix B). Qualitative data was also collected by asking participants for their thoughts and reactions to the CVs and voice samples, which have

been divided into themes and codes. This will be explained further in the next sections. For example, participants would be presented with the Average CV and have to rate the candidate and provide their thoughts about the candidate, based on the CV. The qualitative questions were worded thus, in either English or French: "What stands out for you?" "Reactions to this?" Responses to qualitative questions and debriefings were recorded using Smart Voice Recorder Application (App) on a Nexus 4 device.

After participants completed the CV-based rating and had given their thoughts and reactions for the first-presented CV, the same CV was paired with each of the three voices (FF, FE, and FM) with the same response content. The participants were asked to imagine that each voice was the candidate presented in the written CV in front of them, even though the CV was ultimately paired with each of the three voices. When the first voice was presented, participants were asked to rate the candidate on Likert scales, taking into account *both* the CV and the voice sample. The participant was also asked to provide a verbal reaction to the voice sample in combination with the CV, elicited through prompts similar to the CV-only qualitative questions. The second voice was then presented with the same CV and the candidate rated, followed by the presentation and rating of the third voice. For example, a rater participant would hear FE and see the Average CV, then rate the candidate based on this pairing. The rater would then hear FM and see the same CV, and rate that pairing. Finally, the rater would hear FF and see the same CV and rate the pairing, with response content being the same across all voices.

This procedure, with a CV rated first, followed by ratings of voices paired with the same CV was done for each of the three CVs; raters were therefore presented with a total of three CVs and nine voice samples, counterbalanced for order (see Appendix C for a sample of the counterbalancing procedures).

28

For each rater participant, a short discussion and debriefing was conducted and recorded at the end of the rating session. I shared the motivations and hypotheses for the study, and the participants shared their reactions and thoughts. The debriefing part of the study lasted anywhere from five minutes to half an hour, depending on the participant's availability and interest in the study. The entire rating session took on average approximately thirty minutes, with a minority number of participants taking between forty five minutes to an hour to complete the study because of an extended debriefing conversation.

Data Analysis

Processing and analysis of quantitative data. All of the rating scales were input into Excel (2013) spreadsheets, with different columns and rows for each rater, response content, CV and associated number (1, 2 or 3), accent (FF, FE, or FM) and the rating. The Excel spreadsheets were organized in such a way that each piece of data had its own cell, for manipulation and sorting.

In analyzing data from the rating of the CVs only, a one-way Analysis of Variance (ANOVA) was run with CV-type (Weak, Average, Strong) as the factor. The Brown-Forsythe test for equality of means showed homogeneity of variance for some but not all rating responses. The F-ratio for Items 2, 3, 5, 8, and 10 were not significant at $p = .05$, so responses from these items were not included in further analysis. Post-hoc tests were run with Tukey corrections. When analyzing the CVs and voice samples, separate linear-mixed models were carried out using the participants' ratings for each of the five rated items. In each model, CV type, response content, and speaker were treated as within-groups factors.

Processing and analysis of qualitative data. The participants' qualitative reactions to the materials were recorded and then listened to twice and transcribed verbatim by me, the

author, using the online program Transcribe. To protect the identities of the participants, only I listened to the audio recordings. I have translated all quotations presented below in italics from French into English, and have had my translations verified by a fully bilingual native French speaker.

In analyzing the qualitative data, I created three sets of codes: one for the CVs only (Table 2), one for the CVs and Voice Samples (Table 3), and an abbreviated set of codes for the debriefing conversations. These codes were created when I systematically reviewed the transcribed data, creating codes based on themes extracted from the comments made. The coding itself was checked for reliability by a person who has in-depth knowledge of McGill University's Faculty of Law and the Montreal *stagiaire* recruitment process. The person read the transcripts (which were anonymized) in their entirety. We discussed and then came to a consensus about each coding choice. For the CV-only commentary, 32 codes were created. For the CVs and voice samples, 42 codes were created. For the debriefing conversations, 22 codes were created.

Given the open-ended nature of the questions posed, a wide range of codes were created in order to incorporate the breadth of comments; however, several more general themes emerged which included multiple codes (see Tables 2 and 3). Of the codes created, certain were further given a value of high or low, based on participants' comments (see Appendix D). For example, the code "confident" was coded either with a +, a –, or nothing, meaning the participant thought the speaker sounded confident (+), not confident (-) or the participant simply mentioned confidence without specifying a high or low status.

Table 2
Themes and codes in CV-only qualitative data

Theme	Codes (descriptive titles)
Inferential – inferences about the candidate based on the CV	Flight Risk (potential of leaving soon after hire) Balanced Profile ("fit") What *other* people would think Questions about the candidate (about aspects not included on the CV) Intelligence Likeability Rater questions his or her own reflections (not sure if impressions accurate)
Strength – the relative strength of the CV	Global evaluation (Average, Strong or Weak) Weakness being overcome Interview offer Candidate not working hard enough Comparison with other CV (in the study or in real life)
Content – what is actually listed in the CV	Grades "Academic" CV (a future professor, etc.) Client service experience (work or volunteer) Holes in the timeline Languages Leadership experience Red Flags Research Work experience Non-work/non-academic experience (extracurricular) McGill-specific ("for a McGill student…") Qualities (maturity, etc.)
Procedural – both about the study and the organization's hiring practices	Non-CV specific (comments about CVs in general) Hiring practices Process of the study itself

Table 3

Themes and Codes – CVs and Voice Samples

Theme	Codes (descriptive titles)
Language	Accent
	Taking first language into consideration
	Potential language-related workplace problems
	Language error
	Bilingual
	Non-native
Inferences and opinions about personality from the voice	Assertive
	Cold (unfriendly)
	Confident
	Not polished
	Sincere
	Rater is not impressed or impressed
	Rater is surprised
	Sought-after qualities
	Rater feels solidarity with candidate
	Intelligence
	Fit in with firm culture
Speech style	Tone
	Smooth
	Natural
	Rehearsed
	Well-articulated
	Volume
General comments and comparison	Rater links CV to the voice specifically
	The fit (or not) of CV with voice
	CV better or worse than voice
	Participant compares CV and/or voice with one heard/seen previously
	CV has same effect on the rater regardless of the voice speaking
	Voice has no effect on employability
Content	Good response
	Banal response
	Rehearsed
Procedural – both about the study and the organization's hiring practices	Hiring practices
	Process of the study itself
	Rater unsure about own judgement
	Rater wants more information about voice

Results

Quantitative Data

Results – CVs only. Research Question 1 asks which aspects of a candidate's profile, as demonstrated in their CVs, are valued by recruiters for employability purposes. Using a one-way analysis of variance (ANOVA) on the rating scales for the CVs, several significant differences were found between the Weak, Strong and Average CVs, especially for items relating to academic performance, non-academic experience, and desirability for the position.

First, a main effect was found for perceived intelligence, $F(2,36) = 8.9$, $p = .001$, with Tukey-corrected post-hoc tests showing that the Weak CV was rated less favorably than the Strong CV ($p = .001$) A main effect was also found for candidates being considered strong academically in law $F(2, 36) = 43.01$, $p = .001$, with Tukey-corrected post-hoc tests showing significant differences between all CV types ($p = .001$): the Weak CV was rated lower than the other two CVs, the Strong CV was rated higher than the other two CVs, and the Average CV was rated higher than the Weak CV and lower than the Strong CV. A main effect was also found in comparing each CV's academic performance to that of other McGill Law students $F(2, 36) = 17.05$, $p = .001$. Post-hoc tests with Tukey corrections showed significant differences between all CVs (ps = .001-.024); the Weak CV was rated lower than the other two CVs, the Strong CV was rated higher than the other two CVs, and the Average CV was rated higher than the Weak CV and lower than the Strong CV.

A main effect was revealed for pertinent work experience, $F(2, 36) = 5.5$, $p = .008$. Tukey-corrected post-hoc tests revealed a difference between the Weak and the Strong CV ($p = .007$); the Weak CV was rated less favorably than the Strong CV. A main effect was also found for pertinent extracurricular experience before law studies $F(2,36) = 3.6$, $p = .038$. Post-hoc tests

with Tukey corrections showed that the Average CV was rated higher than the Weak CV ($p=.038$). There was a main effect found for the likelihood of offering the candidate a first interview, $F(2,36) = 5.05$, $p=.012$. Tukey-corrected post-hoc test found a difference between the Strong and the Weak CVs ($p = .010$), with the Strong CV being rated more favorably than the Weak CV. A main effect was seen in the overall rating of the candidate, $F(2, 36) = 12.5$, $p = .001$. Tukey-corrected post-hoc tests found differences between the Average and the Weak CVs ($p = .001$, with the Average being rated more favorably than the Weak CV, and the Weak and the Strong CVs ($p = .034$), with the Strong CV being rated more favorably than the Weak CV. Because of unequal variances, differences between CV types in Items 2, 3, 5, 8, and 10 (in regards to fit, work experience prior to entering law school, hobbies and interests, extracurricular activities during law school, and academic strength prior to law) were not analyzed.

To summarize, the main statistical differences found were between the Strong and the Weak CV, with the Strong CV being rated more favorably than the Weak CV, and some between the Average and the Weak CV, with the Average CV being rated more favorably than the Weak CV. Fewer differences found, mostly in terms of academics, between the Average and the Weak (overall academic strength and comparative academic strength) and the Average and the Strong CVs (overall academic strength and comparative academic strength). The differences have academics in common, which included a gradation of GPA strength. These results will be further analyzed in the qualitative data analysis and discussion sections.

Results – CVs and voice samples. Research Question 2 asks if raters' perceptions of candidates' employability, considering also their CVs and GPAs, is affected by candidates' L2 accent when speaking French. In general, the raters' perceptions were not affected by the candidates L2 accent when speaking French, with the exception of Anglophone voice FE

appearing less confident. The first two of the five rating scale items, rated on a scale of 1 to 6 (with 1 being complete disagreement and 6 being complete agreement) related to candidates' perceived intelligence and confidence, respectively. In regards to intelligence, no main effects emerged, meaning that raters did not differ in their ratings of intelligence as a function of response content, CV, and voice. For confidence, however, a main effect of voice was found, $F(2,49.63) = 5.42$, $p = .007$. Bonferroni comparisons further revealed that the native French speaker (FF) was overall rated higher than the Anglophone (FE) ($p = .008$), and that the Anglophone with a Quebec accent (FM) was rated higher than the Anglophone FE ($p = .047$). This comparison is consistent with the level of French study and language usage each of the speakers using on a daily basis in Quebec. FF is a Francophone who grew up in the French system who uses the language daily, FM uses French at work and at home daily and studied the language at the university level, and FE studied French many years ago as a child, and uses French sporadically in daily life. Therefore, the confidence levels of the speech appear to be perceived as higher as the speaker has studied French and uses the language. The results for the other rating scale items are presented below because other factors interacted with participants' ratings.

Research Question 3 asks whether, when recruiters rate candidates using CVs, GPAs and oral responses, their evaluations are mediated by their judgments of the strength or weakness of candidates' CVs and GPAs. For the third and fourth rating scale items, which focus on the "fit" of the candidate and whether the response adds to the candidacy, main effects were found, as well as significant interactions. For Item 3, regarding the "fit" of the candidate, a significant CV by content interaction was found, $F(4,34.74) = 3.77$, $p = .012$. Post-hoc tests with Bonferroni corrections revealed that this interaction was due to (a) the Strong CV being rated higher than the

Average CV only for Content 1 (QR1) (p = .001) and (b) for the Average CV, Content 2 being rated higher than Content 1 (p = .036). For Item 4, about whether the interview response adds to the candidacy, a significant CV by content interaction was found, $F(4,37.40) = 5.20$, p = .002. Bonferroni-corrected post-hoc tests revealed that this interaction was due to (a) the Average CV being rated higher than the Weak CV only for Content 2 (QR2) (p = .044), and (b) for the Strong CV, Content 1 being rated higher than Content 2 (p = .029). These findings were surprising, as the content for the question responses were pretested to be rated as similar as possible (while using different words), and will be discussed further in the discussion section. Finally, in Item 5, raters rated the likelihood of whether they would offer a first interview to the candidate. A main effect of CV was the only significant effect found, $F(2,34.63) = 6.07$, p =.005. Post-hoc tests with Bonferroni corrections revealed that the Strong CV was rated higher overall than the Weak CV (p = .006). It is of note to mention that there was no difference found between ratings of the Strong CV and the Average CV. This implies that while the Average CV was constructed to be not as compelling as the Strong CV in academic performance and extra-curricular experience, the profile shown on the Average CV was strong enough to the raters that there was no significant effect found between the Strong and Average CVs.

Summary of quantitative results. For RQ1, several differences were found between the Strong, Average and Weak CVs suggesting specific aspects of candidates' on-paper profile which are valued by legal recruiters for employability purposes. Many of the differences, such as academic strength and overall likelihood of interview offer, were between the Strong and the Weak CVs, with fewer differences found, mostly in terms of academics, between the Average and the Weak (overall academic strength and comparative academic strength) and the Average and the Strong CVs (overall academic strength and comparative academic strength). Overall, the

Strong CV was rated more favorably than the Weak CV for likelihood of receiving an interview offer. For RQ2, the accent of the candidates only affected the ratings of the candidates' perceived level of confidence, but not the likelihood of receiving an interview offer. In regards to RQ3, the content of the answer interacted with ratings of perceived "fit," as well as ratings of the response adding to candidacy. Finally, in regards to the likelihood of being offered a first interview, only the Strong CV was rated significantly higher than the Weak CV across all accent types

Qualitative Data

Results – CVs-only

General CVs-only comments. Research Question 1 asks what aspects of a candidate's profile, based on their CVs, are valued by an employer. The participants' comments about the CVs usually reflected the Likert-type rating scales used in the study. In other words, when asked for their reaction to the CVs, many participants made comments further explaining and justifying their numeric ratings. The part of the candidate profile which received the most frequent comments was about academics (strong, weak and neutral comments).

Beyond academics, many participants not only compared and contrasted the three CVs against each other, but also made comments on how the CVs measured up "for a McGill CV" or "in comparison with other McGill students." As well, the participants noted extracurricular involvement, language abilities, work experience (in particular that involving client service and research experience), and the overall strength or weakness of the CV. Table 2 (above) outlines both the themes and codes. An itemization of the coding created for the CV commentary, including a tally of coded instances per CV, is included in Appendix D. Later in this section, commentary about each type of CV (Strong, Average, Weak) will be described.

Some participants balked at the incompleteness of the candidate packages in the study, which were constructed as such for confidentiality and time restrictions. One participant stated:

It's always difficult to tell just based on CVs. I like to have a cover letter because it shows you how they write and how they describe themselves and what they think is interesting about themselves. And the transcripts because it's all great and good to say you have these things but it doesn't mean anything, you have to see the grades.

Therefore, several participants moderated their responses in the absence of some materials which would typically be included in a candidate's application package.

Strong CV. In general, the participants made consistent mention of this CV having strong and impressive law grades. This data supports the categorization of CVs as "Strong, Average and Weak," in terms of academics. For example, the Strong CV and Weak CV received similar numbers of comments about academics which reflected the CV category (i.e., strong academics for the Strong CV and weak academics for the Weak CV), while the Average CV received positive, negative and neutral comments in equivalent numbers. Some participants mentioned that the Strong CV seemed to belong to a "mature" person, specifically because the Masters degree presented in the CV meant the candidate had to be of a certain age.

Several participants mentioned that the CV had an "academic profile," meaning that the research and academic experience was perhaps *overly* present. As one participant said, "This person has the capacity to teach in the Faculty… but probably not to be hired here." Therefore, the focus on academics and perceived weaker client-services experience (regardless of the inclusion of sales experience at a major retailer), seemed to make some participants question the candidate's motivation for applying for a private law firm, in the case of the rater being from

such an environment. Interestingly, this academic element was not noted as being problematic by the one participant who does not work in a private firm.

One of the participants noted that the person studied in Ottawa and worked in Vancouver during the 2010 Olympics. This made the participant question the person's commitment to staying in Montreal, since the general legal recruitment trend is, if all goes well, that the *stagiaires* will stay and become associates in the firms, as training new lawyers is often expensive. If given an interview, this participant would be sure to ask the candidate about commitment to staying in Montreal.

Average CV. Some comments made supported this CV's "average" categorization: "[This CV is] sort of neither here nor there," and "It seems kind of like a standard, middle of the pack, McGill CV." That being said, the participants in the study found a lot to like in the experiences on the Average CV, so much so that quantitative findings showed few significant differences between the Strong and Average CV. Participants' comments reveal what it was about the content in the Average CV that was regarded favorably. The participants noted the CV's client-services experience: "This seems to be a well-rounded person, in the sense that there seems to be more of a mix of work experience… it's not only research focus... but there is also the customer service part." Even though there is customer-service experience included on all of the CVs, the bank experience in the Average CV received relatively more frequent and more positive comments. Some participants also positively noted how "busy," and "active" this candidate seemed. Again, this is in spite of the fact that this CV has a lower word count and slightly less traditionally prestigious activities than the Strong CV, which did not garner such a reaction. For example, the Strong CV includes a Masters degree prior to law school and paid work in a legal environment after the first year of law school (which is rarer and more

prestigious), whereas the Average CV has an undergraduate degree only and work as a research assistant after the first year of law school.

Overall, the reaction from the Average CV was very positive; in fact, the reactions were more positive than was expected for an intended "average" CV. As one participant summarized:

> "*This is the type of candidate that attracts the most out of the three. I find that the person has a variety of more interesting experiences [...] I think that this candidate is much more varied and more complete for me and what I am looking for. The grades are a little less strong than [the Strong CV] but all the same this candidacy is more "complete" for working in a business law firm."*

Therefore, in summary, in terms of academics this candidate was indeed considered "average;" however, for the rest of the profile, this CV was viewed as strong, or even stronger, than the Strong CV, by several of the participants.

Weak CV. Overall, several of the participants made mention of this CV being the weakest of the three. They noticed the relatively low GPA: "The GPA stands out. As being under 3.0. It just does. It's just that number stands out." As well, the first observation another participant made about the CV was that it is the weakest (in the study) and that he or she was, "Pretty sure we wouldn't call to offer them a first interview, compared to the resumes we receive," while another participant observed, "I'm not so hot on this one. The GPA is just not as exciting. And it isn't as if the rest of CV compensates for that." On the other hand, a few of the participants did not make negative comments about the CV at all. One such participant stated, "It's totally an interesting CV, the grades are good, *but they aren't extraordinary, but they are okay…* I think we would still give a first interview."

The other work aspects of the CV, besides the grades, had been manipulated to be less prestigious, but not necessarily significantly less impressive. Overall, the participants seemed less impressed with this CV. The CV presented the candidate's involvement in the "well-being committee" which was generally not seen as a positive. This result will be interpreted further in the discussion section. As well, the hobbies and interests on all of the CVs were created to be more or less equal in terms of including active, solo, and intellectual endeavors. For the most part the participants did not comment on the hobbies on any of the CVs; however, one participant made the comment on the Weak CV about, "The solo travel across Europe to me is not impressive. I don't know why that is in there." This comment further signifies that, overall, this candidate profile was not seen in a positive light, including aspects that would perhaps be viewed more positively on a stronger CV.

Regarding Research Question 1, there are several aspects of a *stagiaire* candidate's profile, as demonstrated in their CVs, which are valued by recruiters for employability purposes, namely: indications of academic excellence, client-facing work or volunteer experiences, and involvement in volunteer and extracurricular activities.

Results – CVs and voice samples. Research Question 2 asks how a recruiter's perception of candidates' employability, considering also the CV and GPA, is affected by an L2 accent when speaking French. Research Question 3 asks, when using CVs, GPAs and oral responses to judge a candidacy, how the raters' evaluations are mediated by the strength or weakness of the CV and GPA. Raters tended to be consistent in evaluating the overall strength and weaknesses of the candidate's paper qualifications (CV), even when CVs were paired with different accents. Appendix D outlines all of the themes and codes for the comments participants made during the part of the study which combined the CVs and voice samples. In each combination, regardless of

the speaker, at least one rater made a comment about the CV seeming just as attractive (or not), regardless of the speaker's accent or speech. For example, one of the participants commented about the Francophone (FF) voice: "[The accent] doesn't change anything... I don't think someone is more intelligent or more confident because they… are Francophone." This participant, after hearing all three of the accents saying the same thing, affirmed that he or she does not think any differently of the paper qualifications based on their different voices. One participant even refused to make differing comments about the three voices, explaining that his or her hiring values are to remain neutral to extraneous factors such as accent and speech style. Another participant provided an opposing thought to voice FF: "My initial instinct is this person is a bit cold… it doesn't do anything to her CV. But it does make me wonder "would we like this person?"" Thus, this rater admitted that the voice was having an effect on her judgement of the person's personality; however, ultimately, she maintained that the CV was still just as strong.

Other participants also did admit that the voice had an effect. One of the participants acknowledged the Anglophone voice FM combined with the Strong CV and QR2 was not what she expected; "I guess I was very surprised. It was not what I would have expected from this person. The lack of polish in their manner of speaking doesn't seem to fit with the person who has accomplished all of these things." As well, another participant had similar comments to make about the same voice, CV and response content combination; "*The impression that this gives is that [the voice] doesn't fit with the CV. The tone, the delivery… the person is mumbling, there is a pronunciation error.*" The Anglophone speaker FM had difficulty correctly pronouncing the word "*s'améliorer*" in response content 2, despite numerous attempts. This may have affected the participants' opinion of this voice. While some participants did mention the voice eliciting

certain judgements, the chance of a student receiving an interview remained fairly consistent

across the CV types (see Table 4).

Table 4
Likelihood of receiving an interview, by CV and voice sample.

CV	Accent	Likelihood of receiving an interview, averaged across raters (scale of 1-6)
CV 1 (Weak)	FE (Anglophone)	4.00
CV 1 (Weak)	FM (Anglophone)	4.31
CV 1 (Weak)	FF (Francophone)	4.46
CV 2 (Strong)	FE (Anglophone)	5.38
CV 2 (Strong)	FM (Anglophone)	4.88
CV 2 (Strong	FF (Francophone)	5.00
CV 3 (Average)	FE (Anglophone)	4.38
CV 3 (Average)	FM (Anglophone)	4.77
CV 3 (Average)	FF (Francophone)	4.70

Qualitative results – summary. Research Question 1 investigates which aspects of the

CV make a candidate employable. The academic record, employment history and student

leadership activities are given the most weight and regard. A strong academic record is looked at

favorably, as is employment experience that is customer-service or research based.

When considering the qualitative data for Research Question 2, the majority of the

participants said that, after hearing the same question repeated by three different voices,

speakers' L2 accents in French (speakers FE and FM) would not affect raters' decisions to offer

or not offer a first interview. As seen in the coded commentary outline (Table 3), these assertions

were made alongside comments made about the firm's hiring practices and culture and

inferences about the voices' perceived personality (in terms of traits such as honesty, sincerity, polish, etc.) Therefore, while some raters said that they don't take the voice into account, the substance of their comments suggest this might not be the entire case. Furthermore, based on the breadth of participant comments about the voices (see Appendix D), it is not the accent of the Anglophone or Francophone speaking that is being judged, but rather the perceived personality behind the voice. This idea will be further explored in the discussion section.

For Research Question 3, when the raters used CVs, GPAs and oral responses to judge the weakness or strength of a speaker's candidacy, their evaluations were mediated by their judgments of the strength or weakness of candidates' CV and GPA. For example, when the CV, including the GPA, was perceived as sufficiently strong, the participants made comments to suggest that they would be interested in meeting the candidate for a first interview. Likewise, when the CV and GPA was perceived as not strong enough, participants made comments to suggest that they would not offer a first interview.

Results based on the hypotheses

The hypotheses were as follows:

<u>Hypothesis 1</u>: Candidates with a strong CV and GPA will be rated favorably regardless of accent.

This hypothesis was supported by both the qualitative and quantitative data.

<u>Hypothesis 2</u>: Candidates with a weak CV and GPA will be rated more negatively – though not overwhelmingly negatively – with an Anglophone accent.

The Weak CV and GPA was rated more negatively regardless of accent, therefore this hypothesis was not fully supported by the data.

<u>Hypothesis 3</u>: Candidates with an average CV and GPA will be rated favorably regardless of accent, though less favorably than candidates with strong CVs and GPAs.

The Average CV candidate was rated favorably regardless of accent, but there was no statistically significant difference between the Strong and Average CVs, thus this hypothesis was not fully supported by the data.

<u>Null hypothesis</u>: No significant relationships will be found between raters' evaluations and any specific type of CV, GPA, or accent.

In regards to the CV and GPA, a significant relationship was found between the Weak and the Strong CV in terms of likelihood of being offered a first interview, with the Strong CV being more likely to receive a first interview than the Weak CV. In terms of accent, the only main effect was found in terms of confidence, with the Francophone FF and the Anglophone FM being rated higher than the Anglophone FE. Therefore, the null hypothesis is supported on some accounts, but not for others.

Discussion

In this study, I investigated whether job-seekers' non-native accents play a role in hiring decisions by simulating the first phase of a particular hiring process. To examine this, job-seekers – Anglophone and Francophone students – identified as having a specific educational background – McGill University, Faculty of Law – were evaluated for specific positions – an entry-level legal position as a *stagiaire.* This study used an experimental design which mixed different accented voice samples and different levels of qualifications, as seen on their CVs. Overall, this study found that neither a native nor a non-native accent played a significant role in candidates' perceived employability, with the exception of a candidate with a non-native accent appearing less confident.

Findings related to Average and Strong CVs. A hypothesis that was not supported was the Average CV being rated any less favorably than the Strong CV, again, regardless of accent. The relatively favourable ratings for the Average CV may have been due to the business-oriented and client-services bank experience listed, as several participants commented that this was pertinent and valuable experience. As well, comments were made about the person linked to the Average CV as being "people-oriented" and "involved in the community." Therefore, while the grades were manipulated to be weaker in the Average CV than the Strong CV, the content relating to work and extracurricular experience was seen quite favorably.

It is also possible that the work and extracurricular experience on the Average CV were regarded *more* favorably than that on the Strong CV. In the participant comments, certain people mentioned that the Strong CV seemed very academic and research-oriented: "This person has the capacity to teach in the faculty… but probably not to be hired here… It looks like the CV of a good student, but not of someone who would be happy here." Even those who seemed to be impressed by the Strong CV tempered their comments: "*It's a good CV, the grades are very good grades for McGill… that being said… the grades and the CV are not everything… you must meet the person… for the fit in question.*" Here, the participant acknowledged that the perceived "academic profile" such as the one on the Strong CV may or may not be a "fit" with their firm culture. This could be due to the fact that the Strong CV was manipulated to have a certain type of degree background, "Now the fact [the degree] is in international affairs ... and human rights... I wonder if there is an interest in business law. Since we're a business law firm." Another participant articulated directly what they liked better about the Average CV than the Strong CV:

> *I think that the profile of this person seems more like someone who could like working in*
>
> *private practice. Just as the other [Strong CV], both seem very intelligent but I think that*

this person is more, a little more business *because of the work experience at the bank, in*

human resources, her role in the student association, volunteering at the Olympic

Games… this is less academic.

Therefore, when the Average CV was evaluated along with the Strong CV, the experience

working at a bank put on the Average CV may have lessened raters' perceptions of difference.

Overall, work experience, especially client-facing or in a business setting, was perceived

favorably by the raters.

  Employability of those who have non-native accents. Another finding was that voice

FE, the more pronounced Anglophone accent, was consistently seen as just as employable as the

other speakers, and the only time this voice was rated less favorably was in regards to

confidence. This finding can be viewed as consistent with Cargile (2000)'s study on "high-

status" Chinese immigrants. Nonstandard speakers of a privileged group, which in this case

would be McGill law students, may not be penalized for being second language speakers.

Cargile's results conform to anecdotal evidence seen by me in my role as Career Development

Officer, by the hiring practices of McGill Anglophone *stagiaires* over the past three annual

recruitment cycles, and what the data suggest, as well. In particular, one participant said,

  "*This, this is a McGill student! Ha! That means that, because, when we compare the*

  candidacies, the McGill ones stand out. That means that McGill students are typically

  stronger than others not just because of the quality of the students but because of the

  experiences that they do, especially the extracurriculars. That adds a lot to a candidacy."

McGill University's Faculty of Law program is selective and therefore considered prestigious by

some; admissions statistics show that the application to admission ratio is about 7 to 1 (Law,

2015) . As well, McGill has the second-top ranked legal university program in Canada,

according to 2015's QS World University rankings. Knowing the students have already passed a judicious and holistic admissions process means that the recruiters may have already had a certain perception of McGill candidates. Thus, having this extra qualitative data about the positive perceptions of McGill candidates from the employers strengthens the connections between this study and Cargile's findings of regarding a high-status group positively. The findings of this study certainly pertain to McGill law students in particular. Given the potential for employers to perceive prestige from the university itself, it is possible that conducting this study with another university or program would garner different results.

The one item on the rating scale that did show a significant effect of accent was in regards to confidence. The Anglophone speaker FE was rated weaker than both Anglophone speaker FM and Francophone speaker FF. Several participants commented on FE's "soft-spoken" and "hushed" voice. Anglophone speaker FE arrived in Quebec less than six months prior to her voice being recorded, and did not use French on a regular basis. In addition to this, she had used her French hardly at all in the past decade, except for one summer program in Quebec seven years prior. The other Anglophone speaker FM had been in Quebec for several years, and spoke French on a daily basis. Therefore, the difference in level of confidence between the two Anglophone speakers is understandable. It is also understandable why the Francophone speaker FF would *not* be rated as more confident than Anglophone FM, given FM's consistent daily use of the French language.

Materials Design – Unintended Findings and Issues

Unintended finding – the perception of "well-being." There was a finding which emerged from four different raters' qualitative evaluations of the Weak CV that was not intended in the original CV design. Some participants raised concern about the Weak CV because the CV

contained activities involving "well-being" (co-chairing a student well-being committee and writing articles about well-being in the student newspaper):

> The fact that she is co-chair of a well-being committee.... *that is super for being balanced but when she starts her practice… in a large firm… it will be very difficult to keep as much…. She must accept that… she must be very available in the office and she will have less time for activities* on the side. *Is this someone who is ready for that, I don't know… maybe yes. But in law practice you must put in a lot of time*.

The intention behind including this content about well-being was simply to simulate a current and realistic McGill CV, as well-being was a popular topic among McGill law students at the time of this study. It became clear, however, when talking to the participants that the topic of well-being and mental health is very charged. Since this study was not created with the aim of exploring this aspect of employability, the only assumption that can be made from this study is that more research is needed to investigate the topic of mental health and well-being within the legal profession, including if and how candidates should disclose their involvements in well-being initiatives in their application packages.

     The response content. Despite rigorous pre-testing with native French speakers who are familiar with McGill law students, quantitative analyses showed interactions of CV by response content in regards to ratings for "fit" and if the response "added" to the candidacy. In general, after hearing the second CV and voice sample combination, the participants did not make any more substantive comments in the qualitative data about the response content, as they generally realized that the study was concerned with the voices themselves, and not what was being said. Thus, I will speculate why the response content had a statistical effect on ratings.

For "fit," the Strong CV was rated higher than the Average CV only in Response Content 1. But for the Average CV, Response Content 2 was rated higher than Response Content 1. The notion of "fit" is a complex, emotional and interpersonal aspect of the hiring process (Rivera, 2015). As per Rivera's recent study on how peoples' emotions can come into play when hiring, perhaps the way the words were used in Responses 1 and 2 emotionally interacted with the participants and with the CVs in a way that is "an art, not a science," (p. 1353).

To further speculate, perhaps Response Content 1 had an effect on the notion of "fit" because this was the only Response Content to use the word "ego" ("*certainly sometimes it* [taking criticism] *is a little difficult on the ego…"*). For the Strong CV, perhaps the raters found that the use of this word "fit" more because of the strength of the CV, and less so for the Average CV. Response Content 1 also includes, "*In my opinion, it [criticism] is a good opportunity to learn"* whereas Response Content 2 states, "*I find it useful and educative to know if my work as good as it can be."* The way Response Content 1 is phrased makes it seem like criticism is a learning opportunity, whereas Response Content 2 is phrased in a way that criticism is useful in terms of knowing where the person stands in terms of employer expectations. Response Content 1 could be seen as slightly more introspective and self-serving, whereas Response Content 2 starts off introspectively, but then the latter part of the phrase could be seen more as the student attempting to make sure the criticism has an actual *impact* on her quality of work. Therefore, for the Average CV, this could been seen as reassuring to the employer that the student would make adjustments to her work based on feedback, whereas for the Strong CV, perhaps the employer would feel less need to have this reassurance. Finally, Response Content 1 opens with *I find that criticism is very important*, whereas Response Content 2 opens with *I am very open to taking*

criticism. The former phrasing is slightly stronger than the latter; such phrasing could have been seen as a better "fit" with the type of person that have the two respective CVs.

Similarly, when rating whether the response adds to a candidacy, this study found that the Average CV was rated higher than the Weak CV only for Response Content 2, and the Strong CV had Response Content 1 rated higher than Response Content 2. Beyond the word choice issues outlined in the previous paragraph, Response Content 2 may be causing an effect for this question perhaps because it was in this voice sample that Anglophone accent FM made a small pronunciation error. While this is response form rather than response content, this pronunciation was noted by some participants. Therefore, it could have had an overall effect on the rating of this content, especially when combined with the different strengths of CVs. No interactions were found at all in regards to Response Content 3, which perhaps means that this response content was not particularly salient for participants in any of the CV-voice combinations. Therefore, in a replication of this study, I would suggest using the wording in Response Content 3 as an exemplar. In a similarly designed study, I would recommend more extensive pilot testing, and perhaps choosing only one version of response content instead of multiple versions.

Overall, many recruiters insisted that it is the *content* of the interview responses that is the most important aspect, not how it is said or with what accent, for example: "I don't think that the accent has any impact whatsoever. I think the vocabulary and the words used maybe have an impact to evaluate the level of comfort in French. But if someone makes a mistake while talking or has an accent is no problem at all." This assertion may be why, despite my best effort to design the study so that the different question response contents are as similar as possible, that the content of the question response did have an impact on certain aspects of the study, namely about "fit" and if the voice sample "added" to the candidacy. The participants in this study were

very focused on the content of both the CVs *and* the content of the voice samples, despite knowing the response content would repeat itself. Ball (1983) intentionally repeated the response content in his study so that the listeners would be "bored" with the content and focus solely on the voices; however, the results in this study would indicate that they recruiters were not bored enough. I suspect this is because their hiring decisions must come from decisions rendered from the content of the employment documents, so it was difficult for the employers to focus on *just* the voice instead of the content. Therefore, the interactions between Response Contents 1 and 2 and CV types means subsequent employability research using multiple versions of job application materials will require more extensive pilot testing, ideally with the same types of raters who will participate in the main study.

 General participant perceptions and the role of context. The participants shared more broad personal insights and perspectives during the debriefing session. Some of the participants did admit that the language and accent was playing some sort of role in the overall evaluation of the candidates: "I think that it is mostly the content of the response… but in making a general impression, [the language] can, maybe, play a *bit* of a role." This recruiter went on to say that while of course someone's level of comfort in a language will affect confidence levels, s/he won't evaluate a second language speaker with the same level of rigour as a first language speaker in that regard. Overall, the comments made by the participants were encouraging to potential candidates, though this may be because of the particular participant pool used.

 The participants in this study seem to be exemplary representatives for linguistically open-minded recruiting. Compared to participants in research from other countries, such as Australia and the United States, (e.g., Seggie et al., 1982; Segrest Purkiss et al., 2006) the participants in this study were much more open and understanding to non-native accents. This is

probably partially due to the bilingual workplaces common in the legal profession in Montreal and partially due to all of the participants knowing both languages at a high level, with some knowing a third language as well. This study itself was often conducted with participants in both English and French, with the participants and myself switching back and forth between the languages as the conversations evolved. The professional environment and participants' language knowledge likely meant that they were conscious and forgiving when it came to hearing non-native French accents, thus enabling them to make more nuanced commentary about the perceived personality behind the voices, instead of simply focussing on the non-nativeness of the candidate. In a city such as Montreal, where a majority of the population has some level of proficiency in an L2, the data collected in this study suggest having multilingual recruiters means more tolerant and accepting hires, at least from a linguistic standpoint.

Non-native French speakers and employment in the legal profession in Quebec: recruiters' perspectives. Overall, the participants were generally open to candidates from both Anglophone and Francophone language backgrounds. No participant said or behaved as if they would prefer candidates who spoke French as a first language. However, everything in a candidate's application had to be considered as a whole, as one participant states: "It's been several years that I have been the director of recruitment here… it's not just the CV and it's not just the voice… it's really the combination of things." Overall, the results of the study should reassure Anglophone students that while their language skills (and in this case, non-native accent) will likely be taken into account, legal recruiters in Montreal seem open to giving first interviews to students whose French allows them to respond to interview questions, no matter their linguistic background or accent.

Future Research

This study is specific to an authentic context of recruitment, using a very targeted and specialized participant pool. Very few past employability studies have used authentic raters and authentic materials to elicit employability ratings, as this study has. Previous research has been much more general in scope, so it is difficult to find points of direct comparison; however, there are factors and conditions in other employability and language research studies that can be used to compare aspects.

Unlike other large employability audit studies (Bertrand & Mullainathan, 2003; Correll et al., 2007), the scale and scope of this study is targeted and specialized. This research targeted a specific subset of employers rather than a general canvas of various employment sectors. Studies aiming at multiple employment sectors may be more generalizable across different contexts than a study such as this, which investigates entry-level employment from candidates of a certain university in the legal profession only; however, this study was created in such a specialized milieu in response to real-world concerns by actual candidates. Large-scale audit studies may apply best for contexts where job-seekers send out a uniform set of application materials en masse to many different types of job postings. The results of such studies don't demonstrate the likelihood of employability for job-seekers who are specific and targeted in their job search and such studies are not useful for a job-seeker looking for work in a specific area or in a specific company.

The results of this study therefore portray important characteristics of the legal employment market in Montreal for McGill University, Faculty of Law students, and can serve as a point of departure for employability research in other specialized job markets and other geographic areas. Future research could include candidates from specific educational

backgrounds, such as various professional academic or vocational training programs. Future research in the legal milieu should also include candidates from various universities program beyond the specific McGill law program. As well, research should be carried out in different fields and industries, as different areas of employment – such as working in construction, in banking, or as an elementary school teacher – likely demonstrate different levels of tolerance for non-native speech.

A further consideration for future research pertains to myself as the researcher. I am an employee of McGill University, and I work in the interest of law students' career development. There is therefore possible that the raters (who were authentic recruiters) were not entirely frank in their commentary and ratings of the candidates. Even though all participants' comments and ratings were kept strictly confidential (and participants were assured of this prior to the study), it was impossible to separate my dual roles as a researcher of Concordia University *and* employee of McGill University from each other. Furthermore, I am an Anglophone myself, researching opinions about Anglophones, so it is possible that some participants may have softened their commentary about Anglophone accentedness. However, the commentary the participants provided seemed forthcoming and honest, so this aspect of the research likely did not dramatically influence the overall results. A replication of this research would ideally use a Francophone researcher who is also not an employee of McGill University; however access to such a specialized pool of raters would prove very difficult for a researcher without any other connection to the participants. The qualitative data and comments gathered from the skilled participants was of a highly nuanced level. Data from participants with this level of expertise is invaluable. Further research should continue to use authentic raters with experience in human resources in their particular field, as opposed to students or random participants.

An additional consideration for future research pertains to the design of the study itself. Due to the very limited availabilities of, and access to, the participants, the materials and procedure were developed so that data collection would be completed in sessions of approximately thirty minutes. The study was designed in this way in order to recruit multiple participants and respect their time, while also gathering statistically significant data. Ideally, future studies will include contain a larger sample of raters, with more CVs and voice samples to allow for factor analysis and other multifactorial tests. Different gendered and accented voice samples would be ideal, as well. In this study, there were three CVs, nine voice samples spoken by three different people, and three different question response contents.

For CVs and question response contents, it is impossible for these materials to be devoid of content and context. They were therefore open to potential contextual bias, depending on how the raters react to the wording choice in the response contents, or the particular experiences on the CVs (i.e. a rater having a particular affiliation with a certain university or place of work). The voices also demonstrate speakers' individual patterns of tone, pitch and speech style, and therefore, again, are open to potential bias based on a listener's personal preference. It is possible that with more versions of each of the materials, the results might not be so affected by one individual speaker. To enact such a study design, participants would need to be available for many hours, or even an entire day of testing. In the legal profession it would be exceedingly difficult for participants to commit to this amount of time, but in other fields, recruiter professionals may be more able to take the time to complete such a study.

There is a final qualification about the overall applicability of the findings. This study was designed to *simulate* a recruiter's first encounter with a candidate with the aim of investigating if an initial interview would be offered based on paper qualifications and a short

voice sample. This study was not an exact replication of the hiring process. While participants did rate the likelihood of a first interview in the rating scale, some participants stipulated that they are only making this judgement based on the materials in the study, but typically they would like to see the cover letter and a complete transcript before making a decision (these materials were omitted in the interest of confidentiality and time constraints). Further still, the *Course aux stages* hiring process is long, lasting from early February through the end of March every year. Therefore, a first interview is not an indication of an ultimate job offer. It is, however, the first step necessary in ultimately securing an offer. Many studies (e.g., McGinnity & Lunn, 2011; Correll et al., 2007) investigate first-interest in an employment scenario, but it is very rare for a study to actually investigate ultimate employment outcomes. In the future, it would be a great contribution to the fields of employability and second language acquisition to conduct a study that investigates employment outcomes in terms of final job offers.

Conclusion

The findings of this study suggest that in the context of a particular recruitment process for entry-level legal positions in the city of Montreal, Anglophone candidates from a particular university are generally not penalized for demonstrating non-native accents of the official language of the jurisdiction, though recruiters may rate some non-native accented candidates' confidence relatively lower than that of other candidates.

This study furthers the state of employability research by addressing several methodological shortcomings. First, the participants are authentic recruiters who hire specific candidates – in this case, McGill University law students. Using specific candidate materials and authentic recruiters has garnered meaningful and applicable, albeit specific, results. Second, the voice samples are mixed with different levels of CVs. In actual hiring practices, first

impressions – such as the first time a candidate's voice is heard and paper qualifications are seen – can be of the utmost importance. The two types of materials – CVs and voice samples – served to better simulate the elements of an employer's first encounter with a candidate. Third, this study tested reactions to Anglophone accentedness when speaking in French. Past research has systematically investigated speakers of a minority language or dialect in contexts where a standard English is typically used. While hiring practices of L2 speakers of English is a larger area of research in the global context, there are still many Anglophones around the world living in non-English areas trying to gain employment. Research should continue to be conducted to help a variety of L1 speakers, in order to help educate as wide a populace as possible when seeking employment in an L2. Finally, given the vast scope possible when researching language discrimination in hiring practices, this study targeted a very specific subset of hiring culture, for the most focussed results possible. These results may or may not be generalizable for other hiring processes, but the findings benefit a specific population – Anglophone McGill Law students – when they endeavor to enter the legal job market in Montreal.

Chapter 3

General Conclusions and Future Work

      Conclusions for the students of McGill University, Faculty of Law. I was inspired to design this study to help McGill Law students better self-assess their own employment profiles, in order for them to make a decision about whether or not to apply for the annual *Course aux stages* Montreal recruitment process. In my capacity as Career Development Officer at McGill University, I was periodically asked by Anglophone students if they should go through the effort of applying to a Montreal firm or organization, given that their native language is not French. The results of this study should serve to reassure McGill Law students that an Anglophone language background, as shown by having a detectable non-native accent in French, does not appear to be a significant factor in a candidate's profile, with the exception of perhaps appearing less confident.

      Based on the results, I have additional recommendations for McGill law students. There are not just one or two factors in a candidacy that can affect employment decisions, as the recruiter said earlier in the manuscript, "It's not just the CV and it's not just the voice… it's really the combination of things." I have seen students erroneously self-exclude from employment opportunities because of one self-assessed weakness. It is important for students to remember that their *entire* candidacy is scrutinized. For example, while grades are a major aspect of the employment profile for an entry-level legal position, each experience on a candidate's CV can give the employer an indication of employability and potential "fit" (Rivera, 2015). Something that may seem not directly related to a legal job, such as working as a bank teller (as on the Average CV), can indeed be seen as relevant experience. In the case of the bank teller job,

many of the study's participants, who work in corporate firms, noted the strength of the CV's client-based experience in a for-profit business environment.

Even something as seemingly pedestrian as working as a waitress or in retail can add to an entry-level employee's skillset. Students should think about what they have developed in the past that can serve the interests of their future employer. To do this, students will need to reflect on the skills they've acquired *beyond* the surface of a particular experience. For example, working with children at a sleep-away camp can demonstrate to a potential employer that the candidate has skills handling difficult clients in an intensive, around-the-clock environment. Finally, students should do extensive research and networking when seeking to apply to an organization, so that they accurately emphasize the aspects of their own experiences that align with the organization's particular culture and needs.

In Montreal, the ability to speak more than one language – especially English and French – is an important facet society, from daily life, to school, to gaining employment. The voices in this study all spoke the same content, which, regardless of accent, signifies that all three of the candidates could accurately articulate a response to an interview question that is the same level of complexity. A candidate without this level of language ability may have a different experience in an interview that is conducted in their second language. That being said, the participant recruiters seemed aware and interested in bilingual McGill students. The recruiters themselves are all L2 speakers of one of the official Canadian languages, and are likely more understanding and conscious employers, in regards to language, than elsewhere in Canada. Employers from across the world could benefit from reading this study and seeing how tolerant these recruiters are, in order to emulate such tolerance in their own practice. As globalization becomes more prevalent, multilingual employees can add a great deal to an organization.

Many of the comments the participants made was more about the perceived personality behind the voice (confidence, warmth, tone, etc.) instead of the language background or accent. As well, confidence is an important factor in a legal recruiting interview and, therefore, I would advise students to practice answering questions about themselves and their CVs in their L2, to enhance others' perceptions of students' confidence when speaking another language. In summary, if a student possesses the level of a French where he or she can accurately articulate the answers to interview questions, then he or she should feel qualified, from a linguistic standpoint, to participate in the *Course aux stages* process.

Future second language and employability research. This study supports contemporary research (e.g., Rivera, 2015) that acknowledges that there is an emotional aspect of the employment and interview process that scholars are grappling with to quantify and account for. As Rivera explains, "In the case of hiring… subjective feelings about job candidates – not just concerns about skills or productivity – can sway the direction of a search," (p. 1342). Because the hiring process is so delicate, and so laden with interpersonal competencies and subjective feelings which are not yet fully explainable, future second language and employability researchers should be very careful when constructing studies to best account for the emotional aspects of recruiting. This is because the emotional aspect of recruiting can become even *more* complicated when the potential subjective feelings a recruiter has about a particular second language are combined with assessing a candidate's profile.

As the literature shows, accent is a charged issue. How someone speaks can suggest to the interviewer everything from where the candidate is from, to what sort of socioeconomic status they could have. In the case of this particular study, Quebec is a society with deeply entrenched issues regarding the use of French and English (Salien, 1998). Therefore, a study that

shows a significant effect of accent in regards to employability really is showing the effect of how the employer interprets a candidates' accent from a sociological, political and geographic standpoint.

There are many studies that need to be conducted to further the field of employability in a second language. Research conducted in Quebec should investigate a variety of employment fields, from a variety of different language backgrounds (Anglophone, Francophone and Allophone). Elsewhere in Canada, the United States, England and Australia, research could be done that is similar to this, using the legal profession or other professions, but with L2 speakers of English. Finally, on a worldwide scale, such research should be expanded from the current trend of investigating L2 English (and now, French speakers) by conducting research with L2 speakers of all other types of languages. In all future research, I would encourage investigators to utilize authentic raters (i.e. recruiters and those who make hiring decisions) whenever possible, for the most realistic results from those who make hiring decisions. I would also encourage researchers to be specific when replicating candidate profiles, such as using those from a certain university or specific employment background. As my research shows, data collected from authentic raters making decisions about realistic and representative candidates can be used to make reasonable assumptions to help actual prospective employees.

In conclusion, because of the complexities of both employability and second language research, similarly focussed studies should be conducted worldwide, but adapting to the varying languages, candidate profiles, and employment sectors of the area of research. Generic canvasses and audit studies (e.g., McGinnity & Lunn, 2011) of entire cities or countries cannot allow researchers to counsel specific types of job-seekers; rather, they can only show broad generalities of the employment climate. Therefore, in order to continue to help those who seek employment,

and in particular those who migrate to areas where another language is spoken, future research should continue to strive to include the specific along with the broad in research designs.

References

Atkins, C. P. (1993). Do employment recruiters discriminate on the basis of nonstandard dialect? *Journal of Employment Counseling, 30*(3), 108-118.

Ball, P. (1983). Stereotypes of Anglo-Saxon and non-Anglo-Saxon accents: Some exploratory Australian studies with the matched guise technique. *Language Sciences, 5*(2), 163-183.

Bertrand, M., & Mullainathan, S. (2003). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination: National Bureau of Economic Research.

Cargile, A. C. (2000). Evaluations of employment suitability: does accent always matter? *Journal of Employment Counseling, 37*(3), 165-177.

Carlson, H. K., & McHenry, M. A. (2006). Effect of accent and dialect on employability. *Journal of Employment Counseling, 43*(2), 70-83.

Clark, E., & Paran, A. (2007). The employability of non-native-speaker teachers of EFL: A UK survey. *System, 35*(4), 407-430.

Cook, V. (1999). Going beyond the native speaker in language teaching. *TESOL quarterly*, 185-209.

Correll, S. J., Benard, S., & Paik, I. (2007). Getting a Job: Is There a Motherhood Penalty? 1. *American journal of sociology, 112*(5), 1297-1339.

Davison, H. K., & Burke, M. J. (2000). Sex discrimination in simulated employment contexts: A meta-analytic investigation. *Journal of Vocational Behavior, 56*(2), 225-248.

Derwing, T. M., & Munro, M. J. (2009). Comprehensibility as a factor in listener interaction preferences: Implications for the workplace. *Canadian Modern Language Review/La Revue canadienne des langues vivantes, 66*(2), 181-202.

Dougherty, T. W., Turban, D. B., & Callender, J. C. (1994). Confirming first impressions in the employment interview: A field study of interviewer behavior. *Journal of Applied Psychology, 79*(5), 659.

Gaies, S. J., & Beebe, J. D. (1991). The Matched-Guise Technique for Measuring Attitudes and Their Implications for Language Education: A Critical Assessment.

Gluszek, A., & Dovidio, J. F. (2010). The way they speak: A social psychological perspective on the stigma of nonnative accents in communication. *Personality and Social Psychology Review, 14*(2), 214-237.

Hellerstein, J. K., & Neumark, D. (2008). Workplace segregation in the United States: Race, ethnicity, and skill. *The Review of Economics and Statistics, 90*(3), 459-477.

Hosoda, M., Nguyen, L. T., & Stone-Romero, E. F. (2012). The effect of Hispanic accents on employment decisions. *Journal of Managerial Psychology, 27*(4), 347-364.

Kalin, R., & Rayko, D. S. (1978). Discrimination in evaluative judgments against foreign-accented job candidates. *Psychological Reports, 43*(3f), 1203-1209.

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.

Laur, E. (2014). Evaluational reactions to spoken French and English in Montreal: Does mother tongue really matter? *The Canadian Journal of Linguistics/La revue canadienne de linguistique, 59*(1), 3-23.

Law, M. F. o. (2015). Admissions Policy. from https://www.mcgill.ca/law-admissions/undergraduates/admissions/policy

Lindemann, S. (2002). Listening with an attitude: A model of native-speaker comprehension of non-native speakers in the United States. *Language in Society, 31*(3), 419-441.

Lindemann, S. (2010). Who's "unintelligible"? The perceiver's role. *Issues in Applied Linguistics, 18*(2).

Lippi-Green, R. (1997). *English with an accent: Language, ideology, and discrimination in the United States*: Psychology Press.

Major, R. C., Fitzmaurice, S. M., Bunta, F., & Balasubramanian, C. (2005). Testing the effects of regional, ethnic, and international dialects of English on listening comprehension. *Language learning, 55*(1), 37-69.

Markley, E. D. (2000). Regional accent discrimination in hiring decisions: A language attitude study.

McGill Faculty of Law: Grade Distribution Table. (2015). from http://www.mcgill.ca/law-studies/policies/grades

McGinnity, F., & Lunn, P. D. (2011). Measuring discrimination facing ethnic minority job applicants: an Irish experiment. *Work, Employment & Society, 25*(4), 693-708.

Medgyes, P. (1992). Native or non-native: who's worth more? *ELT journal, 46*(4), 340-349.

Munro, M. J. (2003). A Primer on Accent Discrimination in the Canadian Context. *TESL Canada Journal, 20*(2), 38-51.

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.

Parton, S. R., Siltanen, S. A., Hosman, L. A., & Langenderfer, J. (2002). Employment interview outcomes and speech style effects. *Journal of Language and Social Psychology, 21*(2), 144-161.

Pingitore, R., Dugoni, B. L., Tindale, R. S., & Spring, B. (1994). Bias against overweight job applicants in a simulated employment interview. *Journal of Applied Psychology, 79*(6), 909.

Purnell, T., Idsardi, W., & Baugh, J. (1999). Perceptual and phonetic experiments on American English dialect identification. *Journal of Language and Social Psychology, 18*(1), 10-30.

Rakić, T., Steffens, M. C., & Mummendey, A. (2011). Blinded by the accent! The minor role of looks in ethnic categorization. *Journal of personality and social psychology, 100*(1), 16.

Rivera, L. A. (2015). Go with Your Gut: Emotion and Evaluation in Job Interviews. *American journal of sociology, 120*(5), 1339-1389. doi: 10.1086/681214

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*(3), 337-353.

Salien, J. (1998). Quebec French: Attitudes and pedagogical perspectives. *The Modern Language Journal, 82*(1), 95-102.

Scassa, T. (1994). Language standards, ethnicity and discrimination. *Canadian Ethnic Studies, 26*(3), 105-121.

Seggie, I., Fulmizi, C., & Stewart, J. (1982). Evaluations of personality traits and employment suitability based on various Australian accents. *Australian Journal of Psychology, 34*(3), 345-357.

Segrest Purkiss, S. L., Perrewé, P. L., Gillespie, T. L., Mayes, B. T., & Ferris, G. R. (2006). Implicit sources of bias in employment interview judgments and decisions. *Organizational Behavior and Human Decision Processes, 101*(2), 152-167.

Seligman, C. R., Tucker, G. R., & Lambert, W. E. (1972). The effects of speech style and other attributes on teachers' attitudes toward pupils. *Language in Society, 1*(01), 131-142.

Appendices

Appendix A

Composition of the CVs

	Strong CV	Average CV	Weak CV
Currently a second-year student at McGill (law)	X	X	X
Undergraduate Degree, Faculty of Arts	X	X	X
Honours Undergraduate degree	X	X	
Undergraduate degree in a location where French *could* be spoken	X	X	X
Law studies commenced immediately after a 4-year degree		X	X
Law studies commenced immediately after graduate degree	X		
Graduate Degree, Faculty of Arts	X		
Undergraduate honours (Dean's list, etc)	X	X	X
Law GPA above 3.0	X	X	
Law GPA above 3.2	X		
Entrance Scholarship – Law	X		
Scholarships, undergraduate	X	X	X
Scholarships, graduate	X		
Legal Information Clinic/Clinic work	X	X	X
Law Journal	X	X	
Student groups, clubs (law)	X	X	X
Summer work after 1st year, law	X (legal environment)	X (research assistant)	X (non-law related)
Work experience – throughout undergrad	X	X	X
Work experience – summers	X	X	X

Leadership (undergraduate)	X	X	X
Leadership – law (student governance, etc.)	X	X	
International experience	X	X	X
French and English	X	X	X
3rd language proficiency	X (fluent)	X (conversational)	
Hobbies and interests	X	X	X
Word count above 400	X	X	
Word count above 430	X		
Total:	**27**	**22**	**15**

Note: X denotes inclusion of the item in the CV.

Appendix B

Procedure

Rater	1
Step 1	Fill out profile and consent form
Step 2	Rate **Weak CV**
Step 3	Fill out CV rating scale
Step 4	Listen to QR 1 – FF
Step 5	Fill out CV+ Voice rating scale
Step 6	Repeat Steps 4 with QR 1 – FM, Step 5 then Step 4 with QR 1 FE, then Step 5
Step 7	Rate **Strong CV**
Step 8	Fill out CV rating scale
Step 9	Listen to QR 2 – FE
Step 10	Fill out CV+ Voice sample rating scale
Step 11	Repeat Steps 9 with QR 2 – FM, Step 10 then Step 9 with QR 2 FF, then Step 10
Step 12	Rate **Average CV**
Step 13	Fill out CV rating scale
Step 14	Listen to QR 3 – FM
Step 15	Fill out CV+ Voice sample rating scale
Step 16	Repeat Steps 14 with QR 3 – FF, Step 15 then Step 14 with QR 3 FE, then Step 15
Step 17	Debrief

Note: **FF** =Francophone accent (Quebecoise); **FE** = French Anglophone accent 1; **FM** = French Anglophone accent 2; **QR** = Response content

Appendix C

Counterbalancing the Materials

	Strong CV			Average CV			Weak CV		
Rater 1	QR3 FE	QR3 FF	QR3 FM	QR1 FF	QR1 FM	QR1 FE	QR2 FM	QR2 FE	QR2 FF
	Average CV			Weak CV			Strong CV		
Rater 2	QR2 FM	QR2 FE	QR2 FF	QR3 FE	QR3 FF	QR3 FM	QR1 FF	QR1 FM	QR1 FE

Note: **FF** =Francophone accent (Quebecoise); **FE** = French Anglophone accent 1;
FM = French Anglophone accent 2; **QR** = Response content; **Rater** = Participant

Appendix D

Explanation and frequency of qualitative codes by CV

Code	Explanation and Commentary	Frequency (positive/negative/neutral)		
		CV - Weak	CV - Strong	CV – Average
ACAD	Academics, GPA (general commentary or averaging comment), + (strong), - (weak or not compelling)	0/9/2	9/1/3	2/2/2
ACADPROFIL	Academic profile (the student is "an academic") in a good way (+) or not (-)	0/0/0	0/0/2	0/2/0
AVER	Commentary about "being average" or middle-of-the-road, or typical, or non-typical (-)	0/0/1	0/1/2	0/0/2
BALANC	The student seems balanced (varied interests), in a good way (+) or in a bad way or in a way that makes the employer worried (-)	1/3/0	1/0/1	1/0/2
CAREER	The student seems focussed on building her career, and is taking steps in that direction	0/1/0	0/0/1	0/0/0
CLIENTSERV	Client services in CV, + (good), - (bad or not enough)	0/0/1	0/0/4	4/0/0
CVNONSPEC	Comment is not specific about the particular CV in question, but about the CVs in general	0/0/2	0/0/0	0/0/1
FLIGHTRIS	The candidate could be a flight risk to the firm	0/0/2	0/0/0	0/0/0
HIRPRAC	Comments made about the firm/organization's hiring practices and culture, in general	0/0/2	0/0/7	0/0/1
HOLE	There is a question about the CV or a question about what the candidate was doing during a certain time. The rater would have liked more information. There is a "gap" or a "hole" in the CV.	0/0/1	0/0/0	0/0/2
INSTI	Institutional commentary, such as, "For a McGill student this looks…" or "McGill students are always…" or "This looks like a typical McGill CV!"	0/0/1	0/0/5	0/0/2
INTEL	Commentary made about the candidates intelligence	0/0/1	0/0/3	0/0/1

INTERVIW	The candidate would be given an interview based on on this CV (+) or would not (-)	1/1/1	0/0/1	0/0/1
LANG	Language knowledge is noted	0/0/0	0/0/1	0/0/1
LEAD	The student exhibits leadership qualities	0/0/0	0/0/2	0/0/1
LIKE	Likeability of the candidate	0/0/2	0/0/0	0/0/2
NONWKNONAC	Non-work and non-academic experience is good (extracurricular) (+) or not good or not enough (-)	3/1/1	3/0/1	7/0/3
NOTWKHARD	Comment made about the candidate not working hard enough, not seeming to be as implicated in activities, etc.	0/0/1	0/0/0	0/0/0
OTH	Comments about what other people would think, not necessarily the rater him/herself	0/0/1	0/0/0	0/0/0
OTHERCAN	Commentary made about another candidate (in the study) + (other candidate is better), - (other candidate is worse)	3/0/1	0/3/1	1/2/0
PROC	Comment about the process and/or procedure of the study (and how it relates to their hiring practices)	0/0/3	0/0/3	0/0/0
PROFIL	Commentary or question about the candidate having a good profile or CV (+) or a questionable profile (-) for their firm	0/5/1	0/1/1	3/2/0
QUAL	The student has sought-after qualities (maturity, etc.) or not (-)	0/0/2	0/1/6	0/0/6
QUEST	The participant has questions about the candidate that aren't being answered in the CV, but this is not necessarily a good or bad thing	0/0/1	0/0/2	0/0/3
REDFLAG	There are red flags in the profile that signify that this student may not be a "good fit"	0/0/4	0/0/0	0/0/0
RESEARCH	Research and writing in CV is noted	0/0/0	0/0/4	0/0/2
SELFQUEST	The participant questions their own reflections, or isn't sure if their reflexes are accurate	0/0/1	0/0/1	0/0/1
STRONG	The student is strong, overall	0/0/0	0/0/4	0/0/1
WEAK	The student is weak, overall	0/0/2	0/0/0	0/0/0
WEAKOV	A weakness in the candidacy is overcome by another aspect ... or not (-)	0/1/1	0/0/0	0/0/0
WORKEX	Work experience is good	3/0/1	1/1/1	2/0/2

Notes: To describe the frequency of the type of comment made, the numbers are listed in order of positive/negative/neutral. For example 1/2/3 would mean there was one positive, two negative, and three neutral comments. "Neutral" comments means that a comment was made as described in the "explanation and commentary" column. Unless otherwise noted, positive comments mean that the rater was being particularly laudatory and negative comments means the rater was being particularly negative.