Assessing contact-induced language change: The use of subject relative markers in Quebec English

Allison V. Lealess and Chelsea T. Smith
University of Ottawa
aleal083@uottawa.ca
csmit066@uottawa.ca

Abstract
To investigate the relationship between language contact and language change, this pilot study undertakes a comparative variationist analysis of relativizers used in subject function in mainstream Canadian and Quebec English. Overall rate of who is the same in both varieties and in multivariate analyses neither variety nor bilingual ability is selected as significant to variant choice, arguing against convergence. Differences in the linguistic conditioning of who across varieties suggest that it is mainstream English, not Quebec English, which has changed. While a breakdown of variant use by sex and age initially suggests contact-induced change, a comparison of speaker cohorts across varieties and a consideration of speaker characteristics fail to confirm convergence. Contributing to the debate on the use of WH-relativizers in spoken English, this study demonstrates that who is a viable option in Canadian English. It also provides evidence against claims that bilingualism is correlated with susceptibility to convergence.

Keywords: language contact; language change; relative markers; Quebec English

1. Introduction
The relationship between language contact and language change is one which has been given considerable attention in the literature. Although it is generally agreed that the lexicon is quite susceptible to change in contact situations, there is debate over whether, or the extent to which, contact can induce change in the grammar of a language. Some, such as Gumperz and Wilson (1971), and Thomason and Kaufman (1988), insist that core syntactic structures or entire grammatical systems may be altered as a result of contact. Likewise, King (2000) reports that because of intensive language contact with English, “core lexical borrowing…resulted in a number of changes in Prince Edward Island French” (176). Thomason (2001: 11) even claims that “all aspects of language structure are subject to transfer from one language to another, given the right mix of social and linguistic circumstances.” However, others maintain that syntactic structures are less susceptible to contact-induced change (Backus 2004) and that cases of...
structural borrowing are uncommon (Winford 2003). Somewhat less controversial is the role that bilingualism is claimed to play in facilitating language change, with, for instance, both Thomason and Kaufman (1988) as well as Heine and Kuteva (2005) remarking on the link between extensive bilingualism and grammatical evolution.

Taking these issues into consideration, the overriding research objective of this pilot study is to assess whether or not language change results from language contact. More specifically, the present study considers whether Quebec English (QcE), in its minority-language guise, has undergone grammatical change as a result of contact with the majority language, Quebec French (QcF).\(^1\) Furthermore, this study aims to address claims that level of bilingualism correlates with susceptibility to change (Heine & Kuteva 2005; Thomason & Kaufman 1988).

In order to investigate these issues, the present analysis adopts the variationist framework (Labov 1963; 1969; Sankoff 1988) and employs as a diagnostic the variable use of relative markers\(^2\) in subject function in restrictive relative clause constructions. In English, it is possible to introduce subject relative clauses by using one of several relativizers, including, \textit{that}, as in (1), WH-forms (i.e. \textit{who}, \textit{which}), shown in (2)-(3), and no marker (\textit{zero}), as in (4).\(^3\)

(1) Like, I have a cousin \textit{that}’s a roadie for a couple of bands right now. (QEC/OW/301:1052)\(^4\)

(2) You could tell who the, you know, actual American citizens were from the Canadians \textit{who} wanna be Americans. (QEC/OW/309:749)

(3) There’s some row housing along there \textit{which} is very, very poor. (QEC/OW/318:73)

(4) It was a neighbour \textit{Ø} came to do the combining for us. (QEC/QC/013:694)

While it has not necessarily been claimed that the relativization system in QcE is undergoing change as a result of contact with QcF, this site is ideally suited to test the susceptibility of a language to convergence for several reasons. First, relativization in subject function in French and English constitutes a true conflict site, an area “at which the structures of the language pair do not match” (Poplack & Meechan 1998: 132). Crucially, it is only at a conflict site where there is the

\(^1\) This pilot study adds to a growing collection of research on this topic in this dialect, including, but not limited to, Adams (2005), Leroux and Jarmasz (2005), Poplack et al. (2006a), Poplack and Walker (2004), Van Herk (2005).

\(^2\) Here, following Tottie and Rey (1997: 225) and Levey (2006: 46), among others, the terms ‘relative marker’ and ‘relativizer’ are used as cover terms for the variants considered in this study.

\(^3\) Note that a handful of other lexical items, such as \textit{what} and \textit{as}, are found to occur in relative marker function in some (non standard) varieties of English, as reported by Tagliamonte, Smith and Lawrence (2005) and by Levey (2006), among others.

\(^4\) Codes refer to city (OW=Oshawa/Whitby; QC=Quebec City), speaker number and line number in the \textit{Quebec English Corpus} (QEC) (Poplack et al. 2006b). Examples are reproduced verbatim from speaker utterances.
potential both for convergence and for evaluating whether it has occurred. In the case of subject relativization, a WH-marker *qui* is the only available option in French, while in Canadian English (CE), WH-forms (*who* and *which*) compete with *that* (D’Arcy & Tagliamonte 2008). Given that this structure could be expected to be targeted on the basis of “typological similarity” (Backus 2005: 323) and “interlingual equivalency” (Johanson 2002: 54; cited in Backus 2005: 323) of the French and English WH-forms, this variable arguably constitutes an area of grammar susceptible to convergence. In fact, change in a language’s relativization system induced by contact has been reported elsewhere: Weinreich (1968: 30) suggests that in American Yiddish, the interrogative pronoun *ver* was extended to relativizing function under the influence of English, as a result of “the identification by bilinguals of Yiddish *ver* with English *who*.” Hill and Hill (1986: 276) contend that while classical Nahuatl lacked true relative pronouns, the Nahuatl interrogative pronoun *tlen* has been extended to relative pronoun function by calquing the Spanish interrogative and relative pronoun *que*. Furthermore, since the historical literature suggests that the introduction of WH-forms into English may have stemmed from Latin and was strengthened under French influence (Mustanoja 1960), it could be argued that QCÉ may be particularly disposed to exploiting these forms as a result of its present-day contact with QCf.

By adopting this variable as a means for detecting contact-induced language change in QCÉ, this line of inquiry opens up a second avenue of research. In particular, while in English WH-forms are widely considered to be the standard relative markers, the actual extent of their use in the spoken language has been subject to much controversy over the last few decades. In particular, though Romaine (1982) contends that “the infiltration of WH into the relative system...has not really affected the spoken language,” Ball (1996) maintains that “the WH-strategy has affected not only standard British English and American English, but non-standard varieties as well.” And while this variable has been given considerable attention in dialects of British English (BrE) (e.g. Ball 1996; Levey 2006; Quirk 1957; Romaine 1982; Tagliamonte 2002b; Tagliamonte et al. 2005)8, and to a lesser extent in American English (AmE) (e.g. Ball 1996; Guy and Bayley 1995), the relativization strategies employed in mainstream CE have been almost entirely overlooked in the variationist literature.9 Given Ball’s (1996: 238) claim that “there is evidence for considerable diversity in spoken English with respect to relativization”, and considering that Walker (2007: 149; drawing

---

5 In particular, French *qui* functions as both a subject relative marker (e.g. *Les filles qui chantaient*) and interrogative pronoun (e.g. *Qui chantaient*?), just as does the English WH-form *who* (e.g. *The girls who sang; Who sang?*). On the other hand, *that* cannot function as an interrogative pronoun in English (e.g.*That sang*?).

6 An indigenous language spoken in Mexico; also referred to as Mexicano.

7 They cite Langacker (1975) and Heath (1972).

8 Note that both Taligamonte (2002b) and Tagliamonte et al. (2005) are studies on northern varieties of British English.

9 There are only two variationist studies known to us which have investigated relative marker use in Canadian English: Poplack et al. (2006a) on relative markers in direct object function in mainstream and Quebec English and D’Arcy and Tagliamonte (2008) on the social conditioning of relative markers in subject and direct object function in Toronto English.
on Brinton and Fee 2001 and on Chambers 1998) acknowledges CE’s characterization as “quintessentially standard,” the use of the WH-forms in this variety is of key interest to the ongoing debate on the status of WH-forms in contemporary speech. Therefore, while here we do not target the CE relativization system in its entirety, a second research goal of this study is to generate a description of one part of the CE relativization system, namely, relative markers as used in subject function, with the additional aims of enabling cross-variety comparison and contributing new information to a contentious debate.

2. Overview of the history of relativization in English
The history of relativization in English can be traced back to the Old English (OE) period, where the demonstrative pronouns se, seo and þœt, the particle þe (either alone or in combination with a demonstrative pronoun) and the zero relative (especially in subject relatives) were all used in relative function (Nevalainen & Raumolin-Brunberg 2002; Tagliamonte 2002b). During the early Middle English (ME) period, the particle þe was lost, being gradually replaced by þat from OE þæt (Romaine 1980), which eventually came to be spelled that (Tagliamonte 2002b). It is also during the ME period that the demonstrative pronouns se, seo and þæt were replaced by that (Levey 2006). In the early ME period then, the apparent tendency was towards simplification of the relativization system by establishing that as the unique relativizer (Romaine 1980).

However, as Romaine (1980:222) notes, during the ME period, this “trend towards simplification was offset...by the introduction of the interrogative pronouns, the ancestors of modern English who, which, etc. in a relative function.” Importantly, the introduction of the interrogative pronouns into the relativization system launched a restructuring in which that, which had been originally used in both restrictive and non-restrictive contexts, came to be confined to restrictive clauses in the late ME period (Romaine 1980; Tagliamonte et al. 2005). Also, while that had been used with both animate and inanimate antecedents, which began to encroach on that’s functional space, being used with both plural and singular animate and inanimate antecedents, as well as with sentential antecedents (Romaine 1980).

Among the WH-forms, who was the last to have developed as a relative marker, with the earliest known example of it in relative function dated 1426 (Rydén 1983). Initially, the distribution of who seems to have been limited to use in formulaic closing expressions in letters (Rydén 1983), and with the Deity as an antecedent (Nevalainen & Raumolin-Brunberg 2002). Eventually though, use of who with other human antecedents “percolated downwards on the scale of personhood,” moving from use with the Deity to use with proper name antecedents and finally to entities lower on the personhood scale, such as collective nouns (Nevalainen & Raumolin-Brunberg 2002: 112). Moreover, the association of human antecedents with who coincided with an increase in the use of that with non-human subjects in the 17th century (Ball 1996). According to Romaine (1980: 224) the introduction of who in the system “had the effect of making English structurally more similar to Latin and French, which for reasons of prestige would have been desirable.”
Finally, while the zero relative dates back to OE (Nevalainen & Raumolin-Brunberg 2002; Tagliamonte 2002b) and continued to occur in the ME period (Nevalainen & Raumolin-Brunberg 2002; Tottie & Harvie 2000), during the Early ME period it is “relatively infrequent” in subject function (Tagliamonte et al. 2005: 77). Whether the decline of zero subject relatives in subject function is related to ambiguity in structures which contain them (Bever & Langendoen 1971) or to the influence of standardisation and the resulting association of zero relatives with informal styles and non-standard speech (Romaine 1982: 78-80; 173) is an issue which “remains controversial” (Tottie & Harvie 2000: 203).

3. Methodology

3.1 Theoretical framework
In order to address our research goals, this study embeds itself within the theoretical framework of Variation Theory (VT; Labov 1963; 1969; Sankoff 1988), which centers on the idea that a particular linguistic function, referred to as the variable, may be fulfilled by two or more competing forms, known as variants (Poplack 2001). The underlying assumption of VT is that variation in speech is neither random nor due to performance errors, but is inherent, and furthermore, that it is constrained, or conditioned, by a variety of contextual factors, linguistic and/or social (Poplack & Tagliamonte 2001). Furthermore, a fundamental aspect of variationist research is its focus on the quantitative analysis of data, which enables the researcher to methodically determine which factors condition the use of different variants, and, more importantly, how these factors condition a variant’s use (i.e. by promoting or discouraging it). Thus, the methodology adopted by variationist research is ideally suited for the purposes of this analysis since it is specifically designed to scientifically uncover and explain the variability which occurs naturally in speech, such as that observed in the relativization system in English. Moreover, by implementing the comparative method outlined in Poplack and Tagliamonte (2001) and in Tagliamonte (2002a), this methodology enables comparisons to be undertaken between, for instance, groups speaking different languages or dialects, and different social groups (e.g. men vs. women) speaking the same language variety.

3.2 Data and methods
The data on which this study is based were extracted from a subsample of speakers (N=16) from the Quebec English Corpus (QEC) (Poplack et al. 2006b), a corpus of oral language data collected via sociolinguistic interviews in Quebec City and Montreal in the province of Quebec, and in Oshawa-Whitby, Ontario, in the early millennium. For our analysis, we extracted data from the interviews of eight speakers each from the Quebec City cohort and from the Oshawa-Whitby

---

10 According to Tottie and Harvie (2000: 202) in the ME period, zero subject relatives were more common than zero object relatives, though the latter “were becoming more frequent;” in Early ME, “zero subject relatives fell into disuse in [Standard English]” although they “survived in the spoken language.” Tagliamonte et al. (2005: 77), citing Bever and Langendoen (1971: 442) and Visser (1963-1973: 538) also indicate that by the Early ME period “zero relatives were more commonly used in nonsubject position.”

11 For specific details on corpus constitution, please consult Poplack et al. (2006b).
cohort;\(^{12}\) the former represents our contact variety of interest and the latter our control variety, or comparative benchmark, against which the possibility of language change is gauged. The subsample was carefully designed so as to have both cohorts equally represented in terms of age (four “younger” speakers (18–35 years old) and four “older” speakers (60+ years old) in each cohort) and sex (four women and four men in each cohort). Furthermore, among the Quebec City cohort, speakers whose assessment of bilingual ability placed them at either end of the bilingualism continuum were of particular interest, and so this cohort was designed to include four speakers considered more bilingual and four speakers considered less, or not at all, bilingual. Stratifying the subsample according to these social factors allows us to address our primary research goals in several ways. First, the younger speakers from Quebec City are presumed to have had a more intense contact with French than the older speakers, since they acquired English after the “Quiet Revolution” of the 1960’s and the passage of Bill 101 in 1977, which “ratified French as the sole and official language” of Quebec (Poplack et al. 2006b). If there has been change in QcE due to contact with QcF, not only should the Quebec City cohort differ from the Oshawa-Whitby cohort, but the change should be most apparent among those speakers having the most contact with French, namely, the younger Quebec City speakers as well as the more bilingual speakers. Likewise, if bilingualism leads to greater susceptibility to language change, then the more bilingual Quebec City speakers should show more signs of convergence than their less bilingual counterparts, regardless of age. Since it is suggested (Klein 1980; Silva-Corvalán 1994) that contact-induced change “may manifest itself in the selection and favouring of one of the competing variants” (Poplack 2008: 193) which has a counterpart in the contact language, and following Adams’ (2005) reasoning, convergence should manifest as a higher rate of WH-use and/or as a levelling of constraints on variant choice (as indicated by a lack of significance of factor groups and/or factor probabilities approaching .5) in the contact variety and among the higher contact speakers as compared to the non-contact and lower contact cohorts.

3.3 The variable context
The variable context was defined as any restrictive relative clause where the relative marker occurred in subject position, as in (1)–(4). Following previous work on this variable by Tagliamonte (2002b) and Tagliamonte et al. (2005), excluded from the analysis were all cases where the relative clause was non restrictive,\(^{13,14}\) where the antecedent was not the subject of the relative clause (e.g.

---

\(^{12}\) The subsample was drawn from the group of speakers which the original investigators (Poplack et al, 2006b) had designated as the ‘core’ group of speakers, who were chosen for the quality of their interview (sound, length etc.), their high or low score on the investigators’ Cumulative French Proficiency Index, and their social characteristics (to achieve a sample with an even distribution of specific social factors).

\(^{13}\) As explained by Tagliamonte (2002b, citing references in Ball (1996)), in contemporary English, non restrictive relative clauses are overwhelmingly marked by WH-forms, so including them in the analysis would drastically inflate the number of WH markers and the resulting analysis would be deceptively skewed.
where the antecedent was the direct or indirect object), where WHIZ deletion occurred (i.e. where “the relative marker and the verb be are deleted” (Tagliamonte et al. 2005:86), as in ‘Sarah married a boy who was born in Hawaii), which had resumptive pronouns, and all tokens which were incomplete, had false starts or were otherwise ambiguous.

Finally, while the literature identifies several other variants such as, what, as and zero, as occurring in various dialects of English, these are not included in the discussion that follows since the zero variant was extremely rare in our data set (only five tokens extracted from three speakers, all of whom are over 60 years old) and no other variants were attested in the speakers sampled here.

In total, 813 cases of subject relatives were extracted: 374 from the control group and 439 from the contact group. Each was coded according to factors which, according to the literature, are hypothesized or found to be relevant to variant choice in English, outlined in §3.4.

### 3.4 Factors

#### 3.4.1 Linguistic factors

##### 3.4.1.1 Type of antecedent

Generally singled out as a key factor in constraining variant choice, an association between who and human antecedents has been reported in both historical and contemporary varieties of English (Ball 1996; Guy & Bayley 1995; Levey 2006; Quirk 1957). Tagliamonte (2002b), however, finds that to be the majority variant with human antecedents in subject position, and, while Tagliamonte et al. (2005) find that who occurs with human antecedents the majority of the time it is used, other variants (that, zero) are more frequent in these contexts. In fact, they observe a greater association between who and the lexical item people than between who and human antecedents. As such, all tokens were coded according to type of antecedent, where we differentiated human antecedents (5), the lexical item people (6), collective nouns referring to humans (7), non-human animate antecedents (8) and inanimate antecedents (9).

(5) The little kid who’s playing Oliver, kinda has his hand over the mic.
(QEC/OW/309:1371)

(6) I know a lot of people that would think you were speaking Greek.
(QEC/OW/311:1370)

(7) I do have a lot of family that live there. (QEC/QC/066:1093)

---

14 Restrictive relative clauses were distinguished from non restrictive relative clauses using semantic considerations. More specifically, only relative clauses which served “to identify their antecedent” (Denison 1998: 278) were classified as restrictive and were therefore included. Cases where the information in the relative clause was non-defining, parenthetical, or supplementary were not included. Likewise, we excluded all cases where the defining nature of the relative clause was uncertain.

15 These are separate environments which show considerable quantitative differences with respect to relativizer use (e.g. Ball 1996; Levey 2006; Quirk 1957; Tagliamonte et al. 2005).

16 Also commonly referred to in the literature as personal antecedents.
We had an old cow that would not let me milk her. (QEC/QC/006:2359)

And you got to open one present that wasn’t from a family member. (QEC/OW/301:205)

3.4.1.2 Sentence type
Sentence type is also reported to constrain use of subject relative markers. In particular, Ball (1996) finds slightly lower frequencies of WH-forms in cleft sentences, and Tagliamonte’s (2002b) results indicate that existential sentences disfavour, and sentence types that are neither cleft nor possessive favour the relativizer who. Levey (2006) reports that who tends to be slightly more favoured in existentials, clefts and possessives than in other types of sentences. In light of these somewhat divergent findings, all tokens were coded for sentence type, where we differentiated cleft (10), existential (11), possessive (12), and other sentence types (13).

It’s a tube that goes up in the arm and round through and down by your heart. (QEC/OW/306:1454)

But there are quite a lot that’s English. (QEC/QC/006:3004)

He had a grin that wasn’t always a grin. (QEC/QC/013:692)

The railway line that went to Lake-Saint-John, they’re wooden rails anyway. (QEC/QC/013:428)

3.4.1.3 Presence of intervening material between the relative marker and the verb of the relative clause
Given Romaine’s (1980: 221) claim that the “WH relativisation strategy appears to have entered the language in the most complex styles” (i.e. in genitive position as opposed to subject position) it could be argued that, in a given environment (e.g. nominative contexts), WH-forms may be used more frequently when the construction is complex. While it is difficult to establish an infallible method for assessing complexity, here we use the distance between the relative marker and the verb of the relative clause as a measure of this, where we distinguish between the presence of intervening elements, as in (14), and the absence of such elements, as in (15).

Note that Levey’s (2006) study is based on the speech of preadolescents.

Of course, there are other measures which may be adopted to assess complexity, including the presence/number of intervening material between the antecedent and the relative marker, whether the relative clause contains nonclausal or clausal arguments (following Tagliamonte et al. 2005), and the length of the relative clause itself (following Tagliamonte et al. 2005, citing Clark & Clark 1977). In their study on complementizer use, Torres Cacoullos and Walker (2009) use adverbial modification, verbal morphology and transitivity to assess syntactic complexity. Such alternatives will be pursued when the speaker sample is expanded beyond this pilot study.
There were some kids that just didn’t have an ear for it. (QEC/OW/309:385)

And you got to open one present that wasn’t from a family member. (QEC/OW/301:205)

3.4.2 Social factors
Several social factors were also considered in the analysis. First, the role played by language variety (contact vs. non-contact) will enable us to determine if there exist differences between the two varieties, which would be a prerequisite in order for us to advance a convergence hypothesis. Second, the contact speakers were coded for their level of bilingual ability (high vs. low/not bilingual), which will further contribute to our primary research goal (since the high bilinguals have had more contact with French) as well as address claims of a link between bilingualism and susceptibility to contact-induced change (Heine & Kuteva 2005; Thomason & Kaufman 1988). Speakers were also coded for age (younger vs. older) for a couple of reasons. First, previous research on this variable has found evidence of a change in time away from WH-forms: Ball (1996:249) reports a decrease in WH-forms in written English in the 20th century, Tagliamonte (2002b) and D’Arcy and Tagliamonte (2008) find that younger speakers disfavour who, and Tagliamonte et al. (2005: 106; citing Baugh 1935: 296) claim that that is “holding its own as the universal relative marker” as WH-forms continue to decrease in frequency. As such, we would expect the non-contact variety, and the contact variety if it is not undergoing contact-induced convergence, to follow this trend away from WH-use. On the other hand, if contact with QcF has led to change in QcE, then QcE, but not mainstream CE, should show evidence of the opposite trend: the younger contact speakers, who have had the most intense exposure to QcF, should have more WH-use than their older counterparts. Finally, given the widely acknowledged role of women in language change, and previous findings that speaker sex is at least mildly implicated in conditioning variant choice in some dialects (Beal & Corrigan 2002; Levey 2006; Tottie & Rey 1997), all speakers were coded for speaker sex.19

4. Results
4.1 General analysis
The overall distribution of variant use by variety is shown in Table 1.

<table>
<thead>
<tr>
<th>Overall distribution of that, who and which in each variety</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-contact English (Oshawa/Whitby)</td>
</tr>
</tbody>
</table>

19 Speakers were also coded according to socio-economic class and level of education. However, due to the limited number of speakers in the sample, the resulting groups were imbalanced, making it impossible to draw any conclusions from the data according to these dimensions. Expansion of the speaker sample beyond this pilot study will enable such social factors to be considered in a future analysis.
In Table 1 we observe that the contact variety completely aligns with non-contact Canadian English in terms of variant distribution: in both varieties, *that* is by far the predominant relativizer,* who* is also fairly common, and variant *which* is extremely infrequent. Together, these findings situate CE somewhere in the middle of the cross-variety continuum of WH-use. More specifically, while *which* is generally found to be rare, the rate of occurrence of *who* varies widely across studies, being extremely frequent in educated Standard BrE (91% according to Quirk (1957)) and rare in older, rural, lower and working class speech (as low as 1.6% according to Macaulay (1991)). Crucially, that variant distribution is essentially identical in both varieties suggests that contact with QcF has not led to an increase in the use of WH-markers in QcE. Note that since *which* is overwhelmingly rare in this data set (N=20), it will not be the focus of further discussion and is excluded from the following analyses. Also, following Ball’s (1996: 233) claim that relative clauses with human antecedents and those with non-human antecedents “represent two different populations” in terms of WH-use, all cases where the antecedent was non-human (N=3) or inanimate (N=196) are excluded from the following analyses. This makes the total number of tokens considered henceforth 592.

Despite the cross-variety similarity in terms of variant distribution, we heed Poplack and Tagliamonte’s (2001: 92) caution against relying solely on variant rates in comparing data sets, and, following their advice, compare the underlying constraints conditioning variant use across varieties. To this end, we undertook two independent multivariate analyses on our two varieties using Goldvarb X (Sankoff, Tagliamonte & Smith 2005). These results are reported in Table 2.

Comparing these results to what has been reported in other dialects of English, we find that, interestingly, it is among the contact speakers where the effect of sentence type most closely parallels that reported by Tagliamonte (2002b) for BrE (existentials disfavour *who*) and by Ball (1996) for AmE (WH-forms are less frequent with cleft sentences.) In terms of antecedent type, CE aligns with what Tagliamonte (2002b) and Tagliamonte et al. (2005) report for BrE, in that variant *that* occurs more frequently with human antecedents than does *who.* Unlike in BrE though, there is no greater association in CE between *who* and *people* than between *who* and human antecedents in either variety (respectively, 42% vs. 47% for the contact variety and 45% vs. 44% for the non-

<table>
<thead>
<tr>
<th></th>
<th>%</th>
<th></th>
<th>%</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><em>that</em></td>
<td>66</td>
<td>245</td>
<td>65</td>
<td>287</td>
</tr>
<tr>
<td><em>who</em></td>
<td>31</td>
<td>116</td>
<td>33</td>
<td>145</td>
</tr>
<tr>
<td><em>which</em></td>
<td>4</td>
<td>13</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>TOTAL</td>
<td>101*</td>
<td>374</td>
<td>100</td>
<td>439</td>
</tr>
</tbody>
</table>

*Due to rounding.

---

20 Variant *who* never occurs in this data set with non-human or inanimate antecedents; only *which* and *that* alternate in this context.
21 Also excluded were two tokens where the antecedent was ambiguous in terms of humanness.
22 This is evidenced by the percentages: for instance, in non-contact English, 44% of all human antecedents occur with *who,* which means that 56% of human antecedents occur with *that.*
contact variety). Note finally that the association between WH-relativization and complexity suggested by Romaine (1980) seems supported here, since the presence of intervening material favours who in both of the CE varieties sampled.

Taken together, we have evidence that QcE shares with mainstream CE English most, but not all, of the constraints on who use. Whether the divergences observed for sentence type are attributable to cross-dialectal differences which exist independent of contact status, or whether they are caused by contact with QcF, is not entirely clear. However, since the divergences do not indicate that a levelling of the constraints on variant choice has taken place in QcE, it is unlikely that they are the result of convergence. Alternatively, since contact English, but not the mainstream variety, shares some of the conditioning reported in other varieties of English, it is possible that the differences may be due to changes having taken place in the non-contact variety.

To further evaluate whether convergence is a possible source of the differences between mainstream and contact CE, we consider whether, and if so

---

Table 2
Results of two independent multivariate analyses of the contribution of linguistic factors to the selection of who in each variety

<table>
<thead>
<tr>
<th></th>
<th>Non-contact English (Oshawa/Whitby)</th>
<th>Contact English (Quebec City)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total N/N:</td>
<td>114/266</td>
<td>145/326</td>
</tr>
<tr>
<td>Corrected mean:</td>
<td>.426</td>
<td>.433</td>
</tr>
<tr>
<td></td>
<td>Prob. % N/N</td>
<td>Prob. % N/N</td>
</tr>
<tr>
<td>Sentence type¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cleft</td>
<td>[.65] 56% 5/9</td>
<td>.17 13% 2/16</td>
</tr>
<tr>
<td>Existential</td>
<td>[.59] 51% 23/45</td>
<td>.21 17% 6/36</td>
</tr>
<tr>
<td>Other</td>
<td>[.48] 41% 67/164</td>
<td>.50 49% 107/219</td>
</tr>
<tr>
<td>Possessive</td>
<td>[.46] 40% 19/47</td>
<td>.61 55% 30/55</td>
</tr>
<tr>
<td>Range</td>
<td></td>
<td>44</td>
</tr>
<tr>
<td>Intervening material²</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>[.54] 50% 6/12</td>
<td>.73 68% 13/19</td>
</tr>
<tr>
<td>No</td>
<td>[.50] 43% 108/253</td>
<td>.49 43% 132/307</td>
</tr>
<tr>
<td>Range</td>
<td></td>
<td>24</td>
</tr>
<tr>
<td>Type of antecedent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human</td>
<td>[.52] 44% 88/199</td>
<td>[.52] 47% 115/244</td>
</tr>
<tr>
<td>people</td>
<td>[.52] 45% 24/53</td>
<td>[.51] 42% 28/67</td>
</tr>
<tr>
<td>Range</td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

¹ Excluded are all tokens where the sentence type was ambiguous (N=1).
² Excluded are all tokens which were ambiguous for this factor group (N=1).

---

23 The numbers in brackets are the results obtained from the first stepping-down run, which is a run where all of the factors are forced into the regression (following Poplack & Tagliamonte 2001: 93). This allows us to undertake a cross-variety comparison of the constraint hierarchies even if the factor group ultimately is not selected as significant.
how, different groups of speaker within each variety are behaving in terms of *who* use. To this end, we undertook two independent variable rule analyses of the contribution of social factors to variant choice, the results of which are reported in Table 3.

### Table 3

Results of two independent multivariate analyses of the contribution of social factors to the selection of *who* in each variety

<table>
<thead>
<tr>
<th></th>
<th>Non-contact English (Oshawa/Whitby)</th>
<th>Contact English (Quebec City)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total N/N:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>114/266</td>
<td>145/326</td>
</tr>
<tr>
<td>Corrected mean:</td>
<td>.427</td>
<td>.423</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Younger</td>
<td>.53 [45%] 68/150</td>
<td>.67 [63%] 117/186</td>
</tr>
<tr>
<td>Older</td>
<td>.46 [40%] 46/116</td>
<td>.28 [20%] 28/112</td>
</tr>
<tr>
<td>Range</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>.55 [48%] 76/159</td>
<td>.37 [28%] 49/175</td>
</tr>
<tr>
<td>Male</td>
<td>.43 [36%] 38/107</td>
<td>.66 [64%] 96/151</td>
</tr>
<tr>
<td>Range</td>
<td>12</td>
<td>29</td>
</tr>
<tr>
<td>Bilingual Ability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low/None</td>
<td>n/a [50%] 92/185</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>.52 [38%] 53/141</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NOTE: When both cohorts are included in a single run, cohort is not selected as significant, and both cohorts have a non-significant input probability of [.50].

A number of important findings emerge from Table 3. First, observe that in the contact variety, level of bilingual ability is not selected as significant to variant choice, a finding which undermines a contact-induced change hypothesis and which fails to substantiate claims by Heine and Kuteva (2005) and Thomason and Kaufman (1988) that bilingualism is linked with greater susceptibility to change. This is further supported by the non-significant data, which suggests that it is the not/less bilingual speakers who may slightly favour *who*; if contact leads to change, then the more bilingual speakers should favour this variant.24 Note also that when both cohorts are included in a single run, variety is not selected as significant to variant choice, again arguing against a convergence hypothesis. On the other hand, in terms of age, while there may be a mild trend in non-contact English for older speakers to disfavour *who* and younger speakers to favour it, in contact English this trend is not only present and strong, but it is exactly consistent with what would be expected if QcF were inducing change in QcE; we return to this observation shortly. Finally, while speaker sex is selected as

---

24 Note however that this effect could be due to a high rate of *who* use among one of the speakers in this cohort, as outlined below.
significant in both varieties, the direction of effect is curiously the exact opposite across varieties, as well as being much stronger in QcE.

Putting all of this all together, we are confronted with a perplexing collection of results: Table 1 indicates congruence between the two varieties in terms of rate, Table 2 reveals cross-variety differences in terms of variant conditioning that are not easily attributable to convergence, and Table 3 provides evidence which both challenges and supports a contact-induced change hypothesis, as well as uncovering an unexpected and peculiar speaker sex effect in QcE. Therefore, to disentangle how differences among social groups may be implicated in these findings, we take a closer look at variant rate among different groups of speakers in each variety. Table 4 reports the overall rate of who by age and sex in contact and non-contact CE.

**Table 4**

<table>
<thead>
<tr>
<th>Overall distribution of who in each cohort by age and sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-contact English (Oshawa/Whitby)</td>
</tr>
<tr>
<td>Younger</td>
</tr>
<tr>
<td>%</td>
</tr>
<tr>
<td><strong>Women</strong></td>
</tr>
<tr>
<td>52</td>
</tr>
<tr>
<td><strong>Men</strong></td>
</tr>
<tr>
<td>36</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
</tr>
</tbody>
</table>

The results shown in Table 4 reveal some interesting trends by age and sex in each variety. In non-contact English, we find that overall, who is the minority variant in both older (40%) and younger (45%) generations, that women use this variant more than their male counterparts regardless of age, and that the men’s use of who remains stable across time. Interestingly, the age effect hinted at in Table 3 seems to be located in the young women of this cohort, who lead in who use at 52%. While this finding is consistent with the widely accepted observations that women use standard variants more than men and that they tend to lead change involving these, note that this finding is opposite the trend of moving away from WH-forms reported in previous research (Ball 1996; Tagliamonte 2002b; Tagliamonte et al. 2005; D’Arcy & Tagliamonte 2008). In the contact variety, while who is the minority variant among older speakers (20%), it is the preferred variant among the younger speakers (63%), exactly what would be predicted if contact with QcF were inducing language change in QcE. However, a closer look reveals that this is largely attributable to the behaviour of young males, who show an overwhelming preference for who at 81%. And even though younger women are using more who than their older counterparts (37% vs. 21% respectively) this is consistent with that observed in the non-contact variety, even if the effect looks stronger. Thus, it is the surprising rate of who use by young men which particularly begs further investigation.

A closer look at the socio-demographic profiles of the two young male QcE speakers provides a partial potential answer for the high rate of who use in this cohort. One of these speakers, who accounts for 2/3 of the all tokens (65%
N=71/110) and for 2/3 of all instances of who (69%; N=62/89) in this group, and whose use of who is a staggering 87% (N=62/71), is classified in the highest socio-economic class, is the only speaker in our subsample to have received a linguistic market indexation score (Sankoff & Laberge 1978) of 4 (the highest possible) and works as an assistant for a lawyer, who happens to be his father. As he is also the less bilingual speaker of the two males in this cohort, his higher use of who is more logically attributable to his socio-economic role and history, than it is to contact induced change. On the other hand, the higher rate of who use (69%, N=62/71) by the other (more bilingual) male in this group is not so easily attributable to details of his socio-economic life, suggesting that contact with QcF may indeed have led to contact-induced change. Still, it would be precarious to generalize this conclusion to the larger QcE population, especially given that among the young QcE women, the less bilingual speaker uses who more than the highly bilingual speaker, opposite a convergence hypothesis. Furthermore, combined with the observed tendency towards greater who use among the young non-contact women, which is suggestive of a change in progress for this variable outside of Quebec, the finding of greater who use in the young contact group fails to conclusively support an argument for contact-induced change in subject relativization in QcE.

5. Discussion and conclusions
We now consider how our findings have addressed our initial research objectives, which were a) to assess whether or not language change results from language contact, b) to determine whether level of bilingualism correlates with susceptibility to change and c) to ascertain the nature of the system of subject relativizers in CE and contribute new information to the contentious debate on the status of WH-forms in spoken English.

In terms of the first objective, we found conflicting evidence regarding whether contact with QcF has led to change in the use of subject relative markers in QcE. First, overall rates of variant use were identical in contact and mainstream CE, and a multivariate analysis did not select either variety or level of bilingual ability as significant to variant choice. On the other hand, rates by speaker age did seem consistent with a contact-induced change hypothesis, although this seemed less likely once other factors were considered. In particular, while young contact speakers use more who than their older counterparts, as a convergence hypothesis would predict, we found evidence that young women in non-contact English also use more who than their older counterparts. And while the young male QcE

25 One anonymous reviewer notes that using age as a social factor could be problematic given “the possibility that WH-use may be a change in progress outside of the contact situation.”

26 While D’Arcy and Tagliamonte (2008) report the opposite trend, namely, that younger speakers disfavour who, their results do not distinguish between young female and young male speakers. And while speaker sex is not selected as significant to variant choice in their analysis, the non significant data indicate a slight favouring effect among women for who use. Therefore, while the results for this study seem to conflict with D’Arcy and Tagliamonte’s (2008) findings, we do not rule out the possibility of a trend among, or incipient change being led by, young women in mainstream CE. Moreover, if this is the case, then the greater use of who among the young contact cohort here could indicate that they are participating in trends/changes affecting mainstream CE,
speakers are leading the way among all cohorts in who use, a closer investigation of the distribution of data and social histories of the speakers in this group suggested that, for at least one speaker, the high use of who is unlikely attributable to convergence. Thus, the age effect in the contact cohort, though suggestive of convergence, cannot be firmly attributed to such an explanation. Also, even if contact with QcF is leading to an increase in the use of who in QcE, it is important to note that it is only affecting the rates of, and not the underlying constraints on, variant use. In particular, while the results of multivariate analyses on the linguistic conditioning of who in both cohorts suggested that QcE does not share with mainstream CE all of the constraints conditioning the variability, comparing variant conditioning in these varieties to other dialects of English (Ball (1996) for AmE; Tagliamonte (2002b) for BrE) suggests that it is the non-contact variety which is slightly divergent from the other varieties. Thus, the dialectal divergences observed in this study are not likely attributable to convergence of QcE with QcF. Moreover, if contact with QcF is promoting who use in QcE, then rate changes are not necessarily accompanied by concomitant changes in variant conditioning. Of course, expansion of the speaker sample beyond this pilot study will permit a detailed analysis of variant conditioning among each age and sex cohort in each variety, to more thoroughly assess the possibility of change, (whether internally or externally motivated) that is suggested by the rate differences reported in Tables 3 and 4.

Looking to our second research objective, our findings fail to substantiate claims made in the literature by Thomason and Kaufman (1988) and by Heine and Kuteva (2005) that bilingualism is linked with a greater susceptibility to change. Not only was level of bilingual ability not selected as significant to variant choice, but the non-significant data suggested that the less bilingual speakers are more likely to use who.

Finally, in line with Ball (1996) but contra Romaine (1980), the results presented here suggest that at least one WH-form, who, has indeed affected spoken English, or at least the varieties considered here, since this variant occurs one-third of the time overall, and more frequently with animate antecedents. Also, the finding that young speakers use this variant more than older speakers suggests, albeit tentatively, that who use may be on the increase. Perhaps the rather healthy incidence of who is not so surprising if we consider Walker’s (2007:149) comment that CE is considered to be “quintessentially standard.” However, given that the methodology (i.e. the Labovian sociolinguistic interview (Labov 1984)) used to collect the language data from which these results are drawn was specifically developed to elicit non-standard, vernacular speech, it is clear that regardless of the “standardness” of CE, who is certainly a viable option for relativizing in subject function in informal, vernacular Canadian English.

which would be consistent with a number of other studies on change in QcE (see Poplack (2008) and references cited therein.) Whether contact with QcF would be facilitating or promoting the adoption or advancement of this trend/change in an issue we leave open for discussion.

27 As one anonymous reviewer notes, education could be implicated here, since “younger speakers may have a higher level of education than older speakers, and education has proven important in other studies of this variable.”
Furthermore, situating the findings reported here to those reported elsewhere on WH-use, this pilot study certainly adds credence to Ball’s (1996: 238) claim of “considerable diversity in spoken English with respect to relativization.”

References


