Second language acquisition and "real" French:
An investigation of subject doubling in
the French of Montreal Anglophones

NAOMI NAGY
University of New Hampshire

HÉLÈNE BLONDEAU
Université d'Ottawa

JULIE AUGER
Indiana University

ABSTRACT

We investigated the French of the first generation of Montreal Anglophones who had had access to French immersion schooling. Our aim was to determine the extent to which these Anglophones had acquired the variable grammar of their Francophone peers and how that was related to the type of French instruction received and to the types of exposure to French. In Montreal French, a subject NP may be "echoed" by a pronoun without emphatic or contrastive effect. Because this is not a feature of standard French, Anglophones who learned French primarily in school were not expected to exhibit it. On the other hand, Anglophones who frequently spent time with Montreal Francophones were expected to have picked it up. To test this hypothesis, we used a database of speech from 29 speakers, varying in their quantity and type of exposure to French. Multivariate analyses determined the degree of correlation of several linguistic and social factors (related to type and quantity of exposure to French) to the presence of a doubled subject. These data were then compared with that for L1 French. Speakers who were more nativelike with respect to the rate of subject doubling and effects of linguistic factors were those who had had more contact with native speakers, especially as adults.

Our goal was to determine the extent to which Montreal Anglophones acquire the variable grammar of their Francophone peers and how that is related to the type of French instruction they have received and the amount of exposure to French they have had. For this study, we examined speech recorded from 29 second

language (L2) speakers of French. The speakers varied along the fluency continuum and in their degree and type of exposure to French at different stages of their lives. They represented the first generation of Anglophones who had had access to French immersion schooling in Montreal. While some had participated in immersion programs, others had attended French schools or "regular" English schools where French was taught as a subject. The speakers were interviewed and recorded in the 1990s by a research team led by Gillian Sankoff and Pierrette Thibault. We compared their acquisition of a particular morphosyntactic pattern to the patterns of use by native (L1) speakers. We also wished to determine which social characteristics of the speakers promoted or impeded successful acquisition of a nativelike pattern.

The pattern we examined is double subject marking, which involves using a strong pronoun + clitic or a noun + clitic. Subject doubling, as shown in (1) and (2), is characteristic of many varieties of colloquial French, including Montreal French (e.g., Auger, 1991; Sankoff, 1982).

(1) *Ma copine elle* venait me voir des fois. (Vincent, p. 14)
    'My friend used to come see me sometimes.'
(2) A ce point là *mes parents ils* en avaient jusque là avec moi aussi. (Joanie, p. 15)
    'At that point, my parents had had it with me too.'

However, because subject doubling is not part of standard French, we would not expect Anglophones who learned French primarily in school to exhibit it.[1] On the other hand, we would expect Anglophones who frequently spoke French with Montreal Francophones to have acquired it. Our study addressed three research questions:

1. Does double subject marking, an L1 pattern not formally taught in school, exist in L2 French?
2. Does the rate of double marking correlate with the type and amount of acquisition and contact with the Francophone community?
3. Does the variation suggest the acquisition of L1 grammar as opposed to the imitation of salient surface structure? That is, could L2 grammar be described by the same linguistic constraints as L1 French?

We found that the answer to each of these questions is "yes." However, this was not a "judgment" or "assessment" study; we did not assume that L1 grammar is necessarily the target of L2 speakers. When L2 speakers differ from L1 speakers, it may be due to incomplete acquisition of L1 patterns or it may result from a desire to express a distinct identity. Given the complex sociopolitical situation in Quebec, this second possibility certainly could not be ruled out. That is, Anglophones who have been required to learn French may not necessarily wish to use this knowledge to blend in with Francophones. Genesee (1987), for example, reported that most Anglophones, even those who do well in their French-immersion classes, do not necessarily seek opportunities to use French outside of the classroom and that their use of French tends to be "reactive" rather than "active." Auger (2002) attributed this reticence to speak French to the fact that Anglo-

phones learning French in school are only exposed to standard Quebec French and thus do not have the linguistic tools that would allow them to communicate comfortably and effectively with L1 speakers. Some consider the absence of Quebecois French features in their speech a handicap and manifest a desire to be more at ease with the local variety (Thibault & Sankoff, 1993).

However, the interviews used in this study also revealed negative attitudes toward Montreal French and/or its speakers, which may explain why some Anglophones refuse, whether consciously or unconsciously, to sound like L1 speakers. For example, Louisa, a fluent L2 speaker who attended a French high school, proudly described the differences between herself and her Francophone classmates in high school, as shown in (3) and (4).

(3) I did not fit in. My accent was different. (tape 33B)
(4) It's just obvious in my speech [that I'm not French-Canadian]. (tape 33B)

In (5) Karen revealed another facet of the complex relationship between the languages.

(5) Elle a fait son: sa maternelle et première année là ⟨d-accord⟩ mais :: mes parents se sont rendu compte qu'elle commençait à parler plutôt en: en joual ⟨ah OK⟩ et ils ont décidé que: ⟨il valait mieux⟩ C'est ça c'est ça ils ont pensé que: que notre français était mieux ⟨OK⟩ et: que c'était mieux qu'elle aille à Ville LaSalle aussi.

'[My sister] did her elementary and first year [at the neighborhood French school] but my parents realized that she was starting to talk "street French" and they decided that—that's it. They thought that [the non-French school siblings'] French was better and that it was better that she also go to Ville LaSalle [an English school with a French immersion program].'

This comment was representative, indicating that L2 speakers knew that the French they learned was not, in most cases, identical to that of L1 speakers.

BACKGROUND

Related research

Rehner, Mougeon, and Nadasdi (1999) provided a thorough review of the literature on the acquisition of French as a second language in Canada. They showed that, for the most part, L2 learners do acquire L1 constraints for a variety of different variables. However, no large-scale quantitative study of subject doubling by L1 speakers was available for comparison to the present study. The following studies were available: a small-scale study of Montreal speakers (Sankoff, 1982), which did not provide quantitative information on the effects of the linguistic variables examined; a study of four Montreal Francophones with some quantitative information (Auger, 1991); a morphosyntactic analysis of pronominal clitics in Montreal French (Auger, 1995); a large-scale quantitative analysis of the speech of Franco-Ontarians (Nadasdi, 1995a, 1995b); a syntactic analysis of subject doubling in New Brunswick Acadian French (Beaulieu & Balcom,

1998); and an analysis of variation in third person pronoun usage (Thibault, 1983). We examined the effects of as many of these linguistic factors as possible in order to compare L1 and L2 speakers.

Previous studies of speakers from this same corpus have analyzed the variable phonological and morphosyntactic patterns that distinguish Montreal French from standard French (Blondeau, Nagy, Sankoff, & Thibault, 2002; Nagy, Moisset, & Sankoff, 1996), the use of discourse markers (Sankoff et al., 1997), and attitudes of speakers (Thibault & Sankoff, 1993) and reactions of community members to different ways of speaking French (Thibault & Sankoff, 1997, 1999). The purpose of these studies was twofold: (i) to describe the use of French by the first generation of Montreal Anglophones (native English speakers) who had had access to French immersion schooling, and (ii) to find correlations between their linguistic patterns and several social factors, including language usage in the social network, language(s) used at work, amount and type of exposure to French in school and in the community, and scholastic level of French.

The present analysis examined previous analyses of sociolinguistic variables to see whether the sociostylistic status of linguistic variables was a relevant dimension in the adoption of (at least some) L1 patterns by Montreal Anglophones. Such comparisons were described extensively in Blondeau et al. (2002). To summarize, Anglophones tend to adopt variants that have no particular stigma associated with them in L1 French. For example, many of these L2 speakers omit the preverbal negator *ne* nearly as frequently as L1 speakers; they also use *on* instead of *nous* as the first person clitic, following a vernacular L1 pattern.[2] In contrast, variants associated with social stratification are generally avoided by Anglophones. Few L2 speakers use the stigmatized apical /r/, and they rarely use compound personal pronouns of the form *eux-autres*, a form that functions as a sociostylistic marker in the L1 community.

Montreal background

As a minority population in a city of French L1 speakers, Anglophone adults have many types of contact with Francophones. Some grew up having regular interactions with Francophone relatives and close friends; others had no French speakers in their social environment as young children but made French friends as adolescents. Still others, who had little or no social contact with Francophones until adulthood, now have a Francophone spouse or partner.

Another important variable is schooling. We included in our sample speakers with three types of school background: traditional French classes in English schools, immersion programs in English schools, and French schools. At one end of the spectrum are subjects who went to French-medium schools; they had the most exposure to French, as they received all of their instruction in French and were surrounded by L1 speakers. At the other end of the spectrum are subjects who were enrolled in traditional programs in English-medium schools; they had the least exposure to French, as they received only a few hours of French-language instruction every week and were surrounded by English-speaking chil-

TABLE 1. *Personal subject pronoun paradigm*

Grammatical Person	Strong Pronoun	Clitic
Singular		
1st person	moi [mwa], [mwe]	je [ʒə)]
2nd person	toi [twa], [twe]	tu [t(y)]
3rd person masculine	lui [lɥi]	il [il], [i], [j]
3rd person feminine	elle [ɛl]	elle [ɛ(l)], [a(l)]
3rd person indefinite	soi [swa]	on [õ(n)]
Plural		
1st person	nous[nu], nous-autres [nuzot]	nous [nu(z)], on [õ(n)]
2nd person	vous [vu], vous-autres [vuzot]	vous [vu(z)]
3rd person masculine	eux [ø], eux-autres [øzot]	ils [i(l)(z)], [j]
3rd person feminine	elles [ɛl], eux-autres [øzot]	elles [ɛ(l)(z)], [i(l)(z)], [j]

Source: adapted from Blondeau (1999:85)

dren. In between these extremes are subjects who were enrolled in French-immersion programs; they received a significant portion of their education in French but were surrounded by other English-speaking children.

Montreal French

Colloquial French differs from standard French in a number of ways. For example, a subject noun phrase or pronoun may be "echoed" by a clitic, without emphatic or contrastive effect. Auger (1995) found that this tendency to double subjects is widely attested in Montreal French. Traditionally, French has been described as a language in which there is one surface marking of the subject, either a pronoun or noun. In its current state, however, it is more accurate to consider it a variably pro-drop language in which the subject may be represented by a noun or (strong) pronoun, optionally followed by a clitic. L1 speakers have carried this process of grammaticalization forward: the clitic is becoming increasingly obligatory (Blondeau, 1999). Table 1 illustrates our usage of the terms "pronoun" and "clitic."

In Montreal French, subject doubling is quite common, but not categorical. Although doubling is obligatory for first and second person pronouns, researchers have disagreed about the status of doubling in third person pronouns. Kayne (1977) and Grevisse and Goosse (1986), discussing standard French, indicated that doubling is either possible but rare (Kayne) or totally impossible (Grevisse & Goosse). Nadasdi (1995b) found that subject doubling is quite common in Ontario French, and Auger (1998) concluded that the alternation between *Eux-autres parlent bien français* and *Eux-autres ils parlent bien français* 'They speak good French' is an integral part of the grammar of colloquial Quebec French. Thibault (1983) argued that, in Montreal French, the only subject pronouns for which doubling is possible are deictics, and that doubling is in fact mandatory. Lexical subjects may be doubled in this variety but do not have to be.

The distribution of doubled and nondoubled subjects for our Anglophone speakers much resembled that for L1 speakers. Although strong pronouns in first and second person must be doubled, lexical and pronominal third person subjects are variably doubled. With third person subjects, doubling may involve a personal pronoun or the demonstrative subject pronoun *ce/ça* (Auger, 1994). For this article, we examine only third person subjects containing a strong pronoun or noun. This envelope of variation is shown in the box in Table 2.[3] When possible, the examples come from our corpus of Anglophone Montreal French (AMF).[4]

The AMF corpus included all instances of doubled subjects preceding the verb. We did not make a distinction between subject doubling and left dislocation[5] for the following reasons. First, some studies have argued for such a distinction, operationalized by criteria such as prosodic separation indicated by a slight pause or absence of liaison (Nadasdi, 1995a), but other studies have shown that these criteria are not always relevant or perceptible (Caroll, 1982; Deshaies, Guilbault, & Paradis, 1992). In addition, L2 speakers may pause and omit liaison for other reasons (e.g., lack of fluency), making the coding of such a distinction unreliable. Finally, Sankoff (1982:85) suggested that the two patterns represent different stages in the process of grammaticalization: "Le clitique deviendrait une composante morphologique du verbe et acquerrait un caractère quasi-obligatoire même lorsque le SN de même fonction est présent dans la proposition" 'The [third person] clitic would become a morphological component of the verb and would acquire quasi-obligatory status, even when a noun phrase with the same referent is present in the clause.'

Evidence that subject clitics in Quebec French already function like a morphological component of the verb comes from the distribution of elements that variably accompany all inflected verbs: (a) strong pronouns are generally doubled by pronominal clitics; (b) lexical subjects are often doubled; (c) quantified subjects, which usually cannot be dislocated, are doubled; (d) for conjoined VPs, subject pronouns must be repeated before each inflected verb; (e) apparent resumptive pronouns are present in subject relative clauses; and (f) a default third person masculine subject pronoun appears in *wh*-questions in which the subject is extracted. This distribution is shown in (6).

(6) a. Toi *(tu) parles bien français.
         'You speak French well.'
     b. Marie (elle) parle bien français.
         'Marie speaks French well.'
     c. Quand quelqu'un (il) se trompe, il faut lui dire.
         'When someone makes a mistake, we have to tell them.'
     d. Je lis des romans, j'écoute de la musique puis je mange des chips.
         'I read novels, I listen to music, and I eat chips.'
     e. Je connais une femme qu'elle écrit des romans.
         'I know a woman who writes novels.'
     f. Qui qu'i a appelé?
         'Who called?'

TABLE 2. *Montreal French sentence subjects*

Strong Subject	Weak Subject	Grammatical Person	
		1st or 2nd	3rd
Noun	Personal clitic	*Liz et moi on va au cinéma.* 'Liz and I go to the movies.' *Liz et moi nous allons au cinéma.* 'Liz and I go to the movies.'	*Mon père il a été une: camionneur* (Jody) 'My father he was a trucker.'
	Demonstrative clitic	(*Liz et moi ce allons…*)	*La culture ici c'est très provincial.* (Brian) 'The culture here it is very provincial.'
	Null	*Liz et moi allons au cinéma.* 'Liz and I go to the movies.' (but *Liz et moi va au cinéma*)	*Mon père a eu une emploi à Ottawa.* (Sandra) 'My father had a job in Ottawa.'
Strong pronoun	Personal clitic	*Nous-autres on était une grande classe.* (Liz) 'We we were a big class.'	*Lui il est beaucoup plus âgé.* (Louisa) 'Him he is much older.'
	Demonstrative clitic	(*Moi ce vais …*)	*Lui c'est mon chum.* (Kelly) 'Him it is my boyfriend.'
	Null	(*Moi vais …*)	*Lui a un appartement.* (Tony) 'Him has an apartment.'
Null	Personal clitic	*J'écoute pour la musique* (Liz) 'I listen for the music.'	*Il n'était pas français* (Liz) 'He wasn't French.'
	Demonstrative clitic	(*Ce vais …*)	*C'est parfait.* (Liz) 'It is perfect.'
	Null	(*Vais …*)	(*Va …*)[a]

[a]There are a few possible null 3rd person forms in Montreal French: sentences beginning with *faut* or *sont*. These were not observed in this corpus.

The only aspect of subject clitics that is difficult to reconcile with viewing them as agreement markers is the fact that doubling with lexical subjects is variable rather than categorical. However, Auger (1998), who discussed variable agreement marking in such languages as Cairene Arabic, English, and even standard French, argued that variable agreement is "neither rare, nor exotic."

Subject doubling in English

Having established that subject doubling is an integral part of Montreal French, let us turn to the patterns of our Anglophone speakers. Possible interference from L1 is often cited when L2 patterns are examined. Therefore, we examined double marking in English as a possible source for any subject doubling that might be found in the French of our L2 speakers. Double marking without emphatic or contrastive intent is rare in English, but it does exist, particularly in southern American dialects (Southard & Muller, 1998; Wolfram & Christian, 1976). Although we found no independent evidence of its existence in Montreal English, (7), (8), and (9) provide examples of doubled subjects in English from several speakers in our corpus. The context indicates that no contrastive or emphatic interpretation characterizes these examples.

(7) *My friend Martine she*'s French. (Greg, tape 2)
(8) *The Quebecois they* know how to party. (Ted, tape 61)
(9) *My sister she*'s a music teacher in Joliet. (Vincent, tape 69)

In order to confirm that a speaker's rate of subject doubling in English did not influence her or his rate of French doubling, material from ten English interviews was examined. (This subsection is adapted from Nagy & Blondeau, 1998.) For each speaker, the first fifteen minutes of the second side of the first tape were used. All instances of doubled subjects were counted, including emphatics, left-dislocations, etc. The total number of clauses produced was counted, and the percentage of doubled subjects was then calculated.[6] Certain clauses that would not permit double marking (e.g., inverted questions, relative clauses, and frozen phrases like 'I don't know' and 'I'm sorry') were excluded. The graph in Figure 1 shows that, for 10 speakers, there was a consistently low (almost nil) rate of doubling in English but widely varying doubling rates in French. Little correlation existed between the percentage of double marking in English and that in French ($r = 0.47$; $p > .05$). The very low rate of doubling in English makes it very doubtful that subject doubling in French could be attributed to transfer from the native language.[7]

METHODOLOGY

The speakers

The speakers studied here were 29 Anglophone natives of the metropolitan Montreal region (17 females, 12 males). Their ages ranged from 20 to 34 at the time
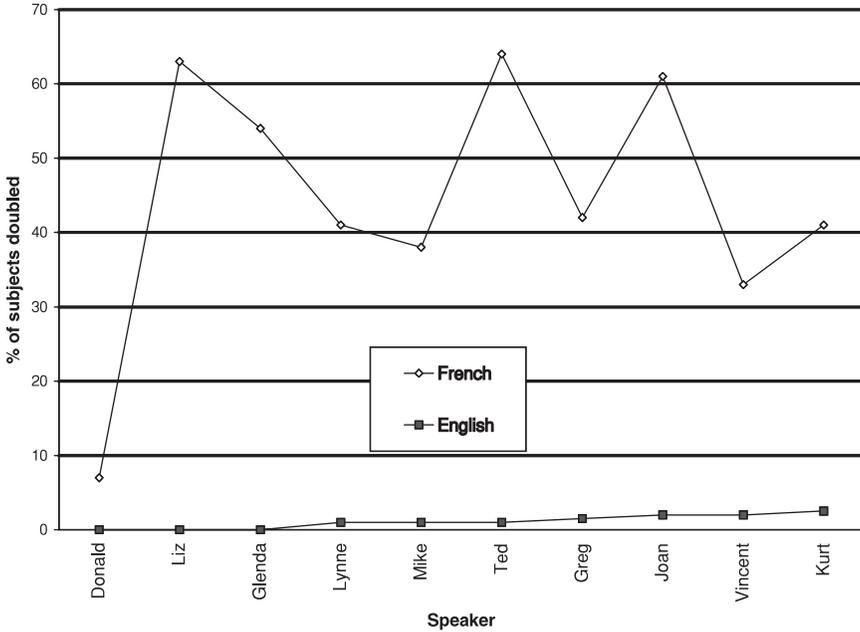
FIGURE 1. Lack of correlation between double marking in L1 and L2.

of the first interview. The speakers were recruited in two ways: one was a self-selecting group of volunteers and the other was a sample of the 1990 graduating class of a high school in the Montreal area. This high school, situated on the border between an Anglophone and a bilingual neighborhood, has received an influx of Francophone speakers over the past several decades. It is an English-language school with two levels of French immersion as well as French taught as a subject (for those less advanced). All the speakers lived in the greater Montreal area at the time of the interview and spoke English with their parents. Nearly all of the speakers came from lower middle class or middle class families. Because of the relative social homogeneity, class was not included as a factor in our study. The speakers did differ with regard to their mode of acquisition of French; the types of exposure they had to French as children; their current degree of contact with Francophones, socially and in the workplace; and the degree to which they used French in their daily lives. Two different scales were developed to operationalize their differential experience with French (Sankoff et al., 1997). We refer to these as the Acquisition and Integration scales.

   The self-selecting participants were recruited through ads seeking bilingual speakers, which were placed in *Voir* and *The Mirror*, two free Montreal newspapers. The high school graduates were located by matching names from the high school yearbook to addresses in the telephone directory. The methodology for interviewing the two groups of participants was the same. They were telephoned

and asked if they would like to participate in an interview about bilingualism in Montreal. The request was made in English; they were then asked if they would be willing to be interviewed in French by a Francophone member of the research group.

Each participant was interviewed in French by a Francophone and, a few weeks later, in English by a non-Montrealer Anglophone. Each interview lasted about one hour. Topics included scholastic and family background; use of French in the workplace; attitudes toward French politics, people, and culture; and incidents where language differences played a significant role. All interviews were tape-recorded. The French interviews were fully transcribed, and relevant material in the English interviews was annotated. The interviews provided two types of data: (i) information about the participants' background and (ii) a record of how they spoke in both French and English.

Linguistic data coding

All clauses with a strong third person subject (pronoun or noun phrase) were extracted from each French-language interview, except for those in the first five pages of the transcript. This resulted in approximately 40 clauses from each speaker, for a total of 1,076 clauses. Nearly 1,000 other clauses had to be eliminated because they had some characteristic requiring categorical doubling (or nondoubling) or they were deemed ambiguous.[8] Each clause was coded for subject doubling: the presence or absence of a subject clitic following the strong subject.[9] This defined the envelope of variation for the study.

Each clause was also coded for 14 linguistic factors, which were grouped as clause type, syntactic, preverbal material, and semantic. These factors were chosen to enable comparison with other research (primarily Nadasdi, 1995b; but also Auger, 1995; Givón, 1976; Sankoff, 1982). The comparisons were made with the understanding that the variable was not defined in exactly the same manner in each of these studies.[10] Because the envelope of variation differed, it would not be appropriate to compare actual values (weights or percentages) across studies. However, the direction of the trends would be comparable. What follows is our selection of factors for each linguistic factor group; each is illustrated with an example.

Subject type

TYPE OF STRONG (FIRST) SUBJECT—DETAILED[11]

    Singular strong pronoun: *lui* 'he', *elle* 'she'

    Plural strong pronoun: *eux-autres* 'they'

    Proper noun: *Madame Pelletier, Paul*

    Common noun: *ma soeur* 'my sister', *les cours* 'the classes'

    Verb or prepositional phrase: *chanter c'est* … 'To sing is …', *de me refuser c'est* … 'To refuse me is …'

    Indefinite pronoun: *quelqu'un* 'someone', *plusieurs* 'several'

TYPE OF STRONG (FIRST) SUBJECT—SIMPLIFIED

    Strong pronoun

    Noun

    Verb or prepositional phrase

TYPE OF CLITIC

il(s)/elle(s) (present or assumed): *Alors eux ils comprenaient* 'So they understood' (Alicia, p. 13)

ce/ça[12] (present or assumed): *Le français c'est pas très difficile* 'French isn't very hard' (Alicia, p. 20)

ambiguous (could be *il(s)/elle(s)* or *ça* or ∅, according to L1 speaker judgment): *Je trouve que les Québécois parlent beaucoup plus vite qu'on a appris à l'école* 'I find that the Quebecois speak much faster than we learned in school' (Alicia, p. 27)

Syntactic

CLAUSE TYPE

Matrix: clause might begin with *et* 'and', *mais* 'but', *ou* 'or'

Subordinate: clause begins with *parce que* 'because', *quand* 'when', etc.

Relative clause: clause begins with *qui* 'who', *que* 'that', or *dont* 'of which'

si clause:[13] clause begins with *si* 'if'

VERB TYPE

Transitive: *parler français* 'to talk French'

Intransitive unergative: *parler beaucoup* 'to talk a lot'

Intransitive unaccusative: *fondre* 'to melt' (as in 'The snow melted')

Passive: *être donnée* 'to be given'

Copula: *être* 'to be'

Preverbal material

RELATIVE CLAUSE

Relative clause modifying subject: *La fille qui est partie parle français* 'the girl who left speaks French'

No relative clause modifying subject: *La fille parle français* 'the girl speaks French'

NEGATION[14]

Preverbal negator *ne* present: *L'école ne m'a pas préparé du tout pour ça* 'School didn't prepare me at all for that' (Brian, p. 13)

No preverbal negator: *Je suis pas français* 'I am not French' (Alicia, p. 14)

PREVERBAL CLITICS

Object clitics, including *en*: *Mes parents ils en avaient jusque là avec moi* 'My parents had had enough of me.' (Joanie, p. 15)

Reflexive clitic: *Les conversations se passent la plupart des fois en anglais* 'Conversations usually take place in English' (Alicia, p. 10)

No clitic: *Tout le monde parle anglais* 'Everyone talks English' (Tammy, p. 20)

OTHER PREVERBAL ELEMENTS[15]

Adverb: *La culture ici c'est très provinciale* 'The culture here is very provincial.' (Brian, p. 13)

Hesitation: *Mais mon père é: était . . .* 'But my father w-was . . .' (Brian, p. 13)

Parenthetical: *Les gens comme mon frère habite dans l'Ouest* 'People like my brother live in the West' (Gloria, p. 6)

Feedback from interviewer: *Mon père parle ⟨hum hum⟩ très bien français* 'My dad ⟨mm-hmm⟩ speaks French very well' (Gloria, p. 8)

PP following head noun: *la fille de ma cousine* 'my cousin's daughter'

None of these: *Tout le monde parle anglais* 'Everyone talks English' (Tammy, p. 20)

Semantic

DEFINITENESS OF SUBJECT

Definite: *la fille de ma cousine* 'my cousin's daughter'

Indefinite, not quantified: *des histoires drôles* 'funny stories'

Quantified:[16] *deux filles* 'two girls', *quelques filles* 'some girls'

Verb or prepositional phrase: *être française* 'to be French'

SPECIFICITY OF SUBJECT
    Specific: *ma mère* 'my mother'
    Nonspecific: *des histoires drôles* 'funny stories'
    Generalizing:[17] *les jeunes d'aujourd'hui* 'young people today'
ANIMACY OF SUBJECT REFERENT
    Animate: *mes amis* 'my friends'
    Material but inanimate: *la table* 'the table'
    Immaterial and inanimate: *ses idées* 'her ideas'
    Place: *notre quartier* 'our neighborhood'

Social data coding

Speakers were coded for 12 social characteristics, which described their language usage, type and amount of formal French acquisition, and type and amount of integration into the Francophone environment. All of this information was self-reported during the course of the interviews. In our sample, there were clear interactions between sex and degree of acquisition and integration.[18] Coding symbols are given for each variant.[19]

Language usage
CURRENT ORIENTATION (in terms of activities, friends)
    1: oriented mainly toward English
    2: oriented toward both communities
    3: oriented mainly toward French
LANGUAGE USED AT WORK
    F: all French
    f: mostly French, some English
    e: mostly English, some French
    E: all English
LANGUAGE USED WITH FRIENDS
    S: has a Francophone partner or roommate with whom French is spoken
    F: has (some) Francophone friends with whom French is spoken
    E: has all Anglophone friends

Type and amount of formal French acquisition
    This scale was based exclusively on the school experience of the speaker and reflected the amount of formal study of French.
ELEMENTARY SCHOOL ACQUISITION SCORE[20]
    1: English elementary school with French as a subject
    2: Partial attendance in French or immersion elementary school
    3: French or immersion elementary school or post-immersion starting in 5th grade
HIGH SCHOOL ACQUISITION SCORE
    Same method as elementary, plus 1 point for secondary education in French.
COLLEGE ACQUISITION SCORE
    Same method as elementary, plus 1 point for post-secondary education in French.
TOTAL ACQUISITION SCORE (combined [summed] score for elementary, high school, and college)
    Combined score of 7 to 8.5: 5 speakers
    Combined score of 5 to 5.5: 10 speakers
    Combined score of 3 to 4.5: 5 speakers
    Combined score of 2 to 2.5: 9 speakers

Type and amount of integration into the Francophone environment[21]

This scale estimated the time spent in a French-speaking environment. It took into account the school experience in terms of exposure to L1-speaking peers as well as extracurricular activities, reported friendships, and family relationships in which French was the language of communication.

CHILDHOOD INTEGRATION SCORE

0: attended English school, studied French as a subject

2: attended immersion school

3: attended French school

+1: participated in French-language activities or had friends or family with whom they spoke French (1 point added for each)

ADOLESCENCE INTEGRATION SCORE

same as above

ADULT INTEGRATION SCORE

same as above

TOTAL INTEGRATION SCORE (combined [summed] score for childhood, adolescence, and adulthood)

Combined score of 7: 4 speakers

Combined score of 4.5 to 6: 8 speakers

Combined score of 3 to 4: 7 speakers

Combined score of 1 to 2: 10 speakers

Method of analysis

Here we examine the effects of each linguistic and social factor on the rate of subject doubling. First, we look briefly at the social factors in univariate analyses, and then we proceed to the multivariate analyses, which allow us to examine several factors simultaneously.

To analyze the data, we employed GOLDVARB 2.0 (Rand & Sankoff, 1990) for Macintosh, a statistical package that uses a logistic regression algorithm to determine the relative effects of each factor. GOLDVARB produces factor weights indicating the strength of each factor: for our purposes, how likely a token would be doubled if it had a particular linguistic or social attribute, independent of its other attributes. Factor weights were calculated using the one-level analysis, and the relative significance of each group was determined using the step-up/step-down analysis.

DATA

Overall, 37% of strong subjects were echoed by a clitic in this sample.[22] The rate ranged from 10% to 83% across speakers. The 29 speakers are characterized in Table 3, where they are listed in order of increasing total Integration score. This table shows that the lowest doubling rates were found with subjects who had the lowest Integration indices, supporting the idea that integration into the Francophone community promotes the acquisition of subject doubling. However, this factor did not account for all individual patterns. Our multivariate analysis enabled us to identify the effect of other linguistic and social factors.

TABLE 3. *Speakers' doubling rates and social characteristics*

| | Dependent Variable | | Independent Variables | | | | | | | | | | | |
| | | | Acquisition | | | | Integration | | | | Current Status | | |
Speaker	Percent Doubled	Total Doubled	Elementary	High school	College	Total	Child	Teen	Adult	Total	Orientation	Work Language	Friends Language
Donald	10	2/21	1	1	0	2	0	0	1	1	1.5	E	E
Jack	14	1/7	1.5	1	0	2.5	0	1	0	1	1	e	E
Jody	67	2/3	1	1	0	2	1	0	1	2	2	f	F
Tony	49	24/49	1	1	0	2	0	0	2	2	2	f	F
Larry	46	12/26	1	1	0.5	2.5	0	0	2	2	2	e	E
Mike	48	10/21	1	1.5	0	2.5	0	1	1	2	2	e	F
Peter	56	9/16	1	1	0.5	2.5	0	0	2	2	1	e	E
Greg	83[a]	5/6	1.5	2	0	3.5	0	1	1	2	2	e	F
Brian	64	16/25	3	1.5	0	4.5	1	0	1	2	3	f	F
Karen	12	7/59	3	2	0	5	1	2	0	3	1	f	E
Glenda	58	7/12	1.5	2	0	3.5	1	1	2	4	1	f	E
Kurt	45	9/20	1	3	0.5	4.5	0	2	2	4	2	f	E
Janie	57	4/7	2.5	2.5	0	5	2	1	1	4	1	e	E

Louisa	56	18/32	2	1	2	5	1	2	1	4	1	f	E
Kelly	48	16/33	3	2.5	0	5.5	2	1	1	4	1	F	E
Spencer	49	22/45	3	3	0.5	6.5	1	1	2	4	3	f	S
Ross	40	6/15	2	1	0	3	2	1	1.5	4.5	2	f	S
Joan	70	33/47	1	1	0	2	1	1	3	5	2	e	S
Gloria	21	5/24	3	1	1	5	3	1	1	5	1	e	F
Janice	37	11/30	2.5	2.5	1.5	6.5	2	2	1	5	1	e	E
Lynne	52	42/81	2	2.5	1	5.5	2	2	1.5	5.5	2	f	F
Alicia	28	8/29	2.5	2.5	0	5	2.5	2	1.5	6	1	f	F
Sandra	50	2/4	3	2	0	5	3	2	1	6	1	f	E
Tammy	20	10/50	2	3	0	5	2	2	2	6	3	F	S
Joanie	58	42/73	3	2.5	0	5.5	3	2	1	6	1	e	F
Jocelyn	41	7/17	1	1	0	2	3	2	2	7	1	f	E
Ted	67	42/63	3	3	1	7	3	2	2	7	3	f	S
Liz	42	10/24	3	3	2	8	3	3	1	7	1	f	F
Vincent	46	23/50	3	3	2.5	8.5	2.5	2.5	2	7	2	e	S

[a]Greg's high doubling rate might have been due to a small number of tokens and the possible misinterpretation of some hesitations and repetitions as doubling. It is his low degree of integration with French speakers that makes us suspect this. However, it is interesting to note that Greg provided one of the rare examples of subject doubling in English (see (7)).

ANALYSIS

Linguistic analysis

Let us first consider the overall effects of the linguistic factors. We compared their effects in our AMF corpus to the significant correlation effects that other researchers reported for L1 speech. The L1 data for all factors were taken from Nadasdi (1995a, 1995b), which also provided an in-depth discussion of the reasons for the rankings. Sankoff (1982) supplied the concurring data for clause type and definiteness. Table 4 compares the ranking of factors for L1 and L2 data. The "L1 hierarchy" column shows the ranking of the factors in L1 speech; the factor correlating to the highest doubling rate within each factor group is listed first. The "L2 hierarchy" column lists the factors we examined, ranked from highest to lowest weight. The last three columns present the data for our L2 sample, given as factor weights, percentages, and frequencies. Our analysis here focuses on the weights. The application value is "doubled": larger values mean more doubled subjects. Where the L1 and L2 hierarchies agree, L1 = L2 appears. Where most parts of the ranking are the same but the order of two factors are reversed, L1 ≈ L2 appears. The present analysis includes 889 tokens, the total number of tokens after "knock-outs."

In the aggregate, L2 speakers shared the grammar of L1 speakers to a great extent. For all but two factor groups, where we have L1 data and significant effects for L2, the factors are ranked in the same order. In the case of clause type, we have only four tokens in relative clauses, and the fact that the ranking of subordinate and relative clauses is out of order is of minimal interest. The other factor where the L1 and L2 rankings do not match is other preverbal elements. The L1 data compares preverbal hesitations versus adverbs. In L2 speech, hesitations often serve a different purpose than in L1 speech. In all likelihood, then, we are not comparing the same types of hesitations. Alternatively, the differences in the two coding systems may play a role. Nadasdi was comparing hesitations and adverbs, whereas we considered more categories.

The significant linguistic factor groups are listed here in descending order of effect on the L2 doubling rate: clitic type, subject type (detailed), other preverbal elements, specificity, preverbal negator, clause type, definiteness.[23] Like L1 speakers, our L2 speakers favored doubling when the clitic is *ce*/*ça* rather than a personal pronoun. They were also more likely to double subject pronouns than any other type of lexical NP, and they were more likely to double proper nouns than common nouns. When the head of the subject is separated from the verb, the rate of doubling depends on what sort of material intervenes. When there is a preverbal hesitation, parenthetical, interruption by the interviewer, adverb, or prepositional phrase, doubling is favored. When none of these occur, it is disfavored. Again, this parallels the behavior of L1 speakers.

With respect to semantic factors, specificity and definiteness both favor subject doubling, whereas quantified, indefinite, generalizing, and nonspecific subjects disfavor doubling. In this respect, it is striking to note that L2 speakers behaved like their L1 counterparts, and that these same factors have been found

to favor doubling and agreement marking in such diverse languages as Cairene Arabic, Zulu, Hungarian, colloquial English, and Porteño Spanish (Auger, 1998).

When there is a preverbal negator, *ne*, doubling is strongly disfavored. Indeed, *ne* is most often omitted in spontaneous Montreal French (Sankoff & Vincent, 1977) and its presence is associated with formal registers and topics, and yet subject doubling is characteristic of informal settings and colloquial speech (Sankoff, 1982). Our L2 speakers appeared to match that pattern: doubling is a vernacular feature and thus rarely co-occurs with the nonvernacular *ne*.[24] The presence of an intervening relative clause or nonsubject clitic does not have a significant effect in the L2 data.

Finally, like L1 speakers, L2 speakers doubled their subjects more often in main clauses than in subordinate clauses. As noted earlier, the apparent exception to the effect on doubling of a relative versus a nonrelative subordinate clause is likely due to the small number of relative clauses contained in our corpus.

Several factor groups that are significant for L1 speakers are not significant for L2 speakers: animacy, verb type, relative clause presence, and preverbal clitic. These differences, listed in Table 5, may indicate slight differences between the L1 and L2 grammars or methodological differences.[25] Overall, however, a great deal of similarity exists in the hierarchy of factors pertinent to the two grammars. In other words, L2 speakers' doubling rate is affected in much the same way by the linguistic factors as that of L1 speakers.[26]

Social factor analysis

So far, we have examined the data for all speakers combined. However, speakers whose doubling rates greatly differ (ranging from 10% to 83%) may be operating with different grammars. We now turn to the analysis of the variation among the speakers. We explored interspeaker variation by conducting further analyses in which the linguistic factors found to be significant were retained and different social factors were considered in the multivariate analysis.

The selected social factors split our speaker sample in different ways. The need to examine all of these arose from the lack of correlation among them. This is easily seen in the graph in Figure 2. The speakers are arranged along the x-axis in order of increasingly positive orientation toward French. Their scores for six social factors are indicated by the thin lines. The percentage of subjects doubled by each speaker is indicated by the thick line.[27] A quick glance reveals the lack of correlation among any of these factors.

Univariate analyses.   To confirm the lack of effect of any one social factor, Pearson correlation tests were conducted. The correlation coefficient *r* was calculated for the doubling rates of the speakers and their scores for the social factors. As Table 6 shows, none of the values of *r* are very close to either -1 or 1. Thus we can discount the idea that any one of these factors is solely responsible for the differing doubling rates of the speakers. None of the correlations reach significance, and so very little variation is accounted for.

TABLE 4. *Comparison of linguistic factors in L1 and L2 French (*N = 889, *Input value* = 0.447*)*

Group	L1 Hierarchy		L2 Hierarchy	Weight	% Doubled	Frequency
Syntactic						
Subject type	Pronoun >	L1 = L2	VP/PP	0.904	95	18/19
	Proper N >		3sg. f. pronoun	0.886	79	23/29
	Common N		3sg. m. pronoun	0.817	68	40/59
			3pl. m. pronoun	0.653	48	16/33
			Proper noun	0.620	63	66/104
			Common noun	0.400	38	242/645
Clause type	Main >	L1 ≈ L2	Main	0.535	47	359/759
	Subordinate >		Relative clause	0.444	50	2/4
	Relative		Subordinate	0.372	39	42/109
			*Si*clause	0.058	12	2/17
Clitic type	*Ce/Ça* >	L1 = L2	*Ce/Ça*	0.918	86	222/259
	Personal		Personal	0.271	29	183/630

Preverbal material						
Negation	Ø >	L1 = L2	Ø	0.517	46	403/869
	ne		*ne*	0.046	10	2/20
Other elements	Adverb >	L1 ≠ L2	Hesitation	0.786	61	30/49
	Hesitation		Parenthetical	0.782	75	15/20
			Back-channel feedback	0.759	64	14/22
			Adverb	0.750	66	25/38
			PP	0.631	55	31/56
			None	0.434	41	290/704
Semantic						
Definiteness[a]	Definite >	L1 = L2	Definite	0.531	49	355/725
	Quantified >		Quantified	0.362	20	24/119
	Indefinite		Indefinite	0.308	33	9/27
			Not applicable	0.479	94	17/18
Specificity[b]	Specific >	L1 = L2	Specific	0.549	48	270/558
	Nonspecific		Nonspecific	0.427	24	17/70
			Generalizing	0.416	45	118/261

[a]Significance is unclear from the step-up/step-down analysis: neither selected nor eliminated.
[b]Significance is unclear from the step-up/step-down analysis: both selected and eliminated.

TABLE 5. *Nonsignificant linguistic factors in L2 (significant in L1)*

Group	Variant	Weight	Percent Doubled	Total Doubled
Animacy	material, inanimate	0.636	59	13/22
	inanimate	0.560	77	138/180
	animate	0.494	34	219/642
	place	0.295	76	35/46
Verb type	transitive	0.600	36	88/242
	intransitive (unergative)	0.531	31	30/98
	passive	0.461	30	3/10
	être	0.450	57	264/460
	intransitive (unaccusative)	0.446	25	20/79
Relative clause in subject	relative clause	0.522	61	27/44
	none	0.499	45	378/845
Preverbal clitics	*en*	0.672	50	1/2
	none	0.504	46	384/826
	reflexive	0.482	31	5/16
	object clitic	0.429	33	15/45

Multivariate analyses. Next we examine the effects of each social variable. Because of the lack of correlation between doubling rate and any one linguistic or social factor, we turned to a more sophisticated method: multivariate analysis. Let us illustrate the sort of problem that this can resolve. Less integrated L2 speakers used the preverbal negative marker *ne* more than did more integrated speakers, who adapted to its virtually categorical absence in L1 speech (Sankoff & Vincent, 1977). We saw before that the presence of *ne* correlates with a decrease in the doubling rate. But a question remains: is the rate lower in the sentences with *ne* because of the linguistic factor (negation) or directly because of the social factor (integration)? Multivariate analysis allowed us to tease apart such complications.

Our multivariate analysis contains the integration and acquisition variables for three life stages, language used at work, language of friends, current orientation, and sex of the speakers, as well as the significant linguistic factors. Only language used at work and adult integration correlate significantly to the doubling rate: that is, more exposure to French favors doubling. Table 7 presents these factor weights. We attribute the lack of significance of current orientation and language of friends to insufficient data in certain cells.

This analysis suggests that it is the speakers' recent environment that most influences this linguistic variable. The language used at work and the degree of integration into the Francophone community as adults correlate significantly with the doubling rate. The near categorical lack of doubling among speakers who never speak French at work or in general highlights the vernacular quality of the variable. Subject doubling is only really acquired by people who actually speak French with Francophones.
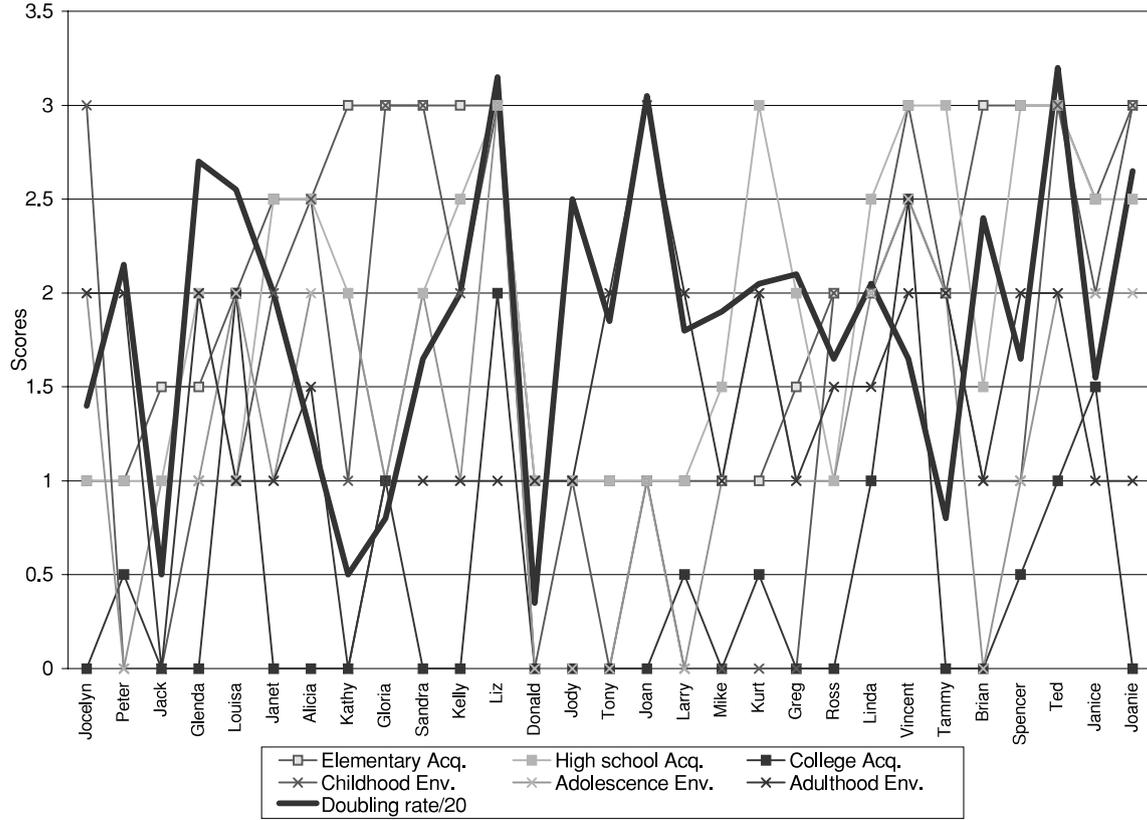
FIGURE 2. Noncorrelation of social factors and doubling rate.

TABLE 6. *Correlation tests for social factors*
and doubling rate

Social Factor	r	P^a
Elementary acquisition	0.03	0%
High school acquisition	0.17	3%
College acquisition	0.23	5%
Total acquisition	0.18	3%
Childhood integration	0.13	2%
Adolescence integration	0.09	1%
Adult integration	0.38	14%
Total integration	0.25	6%
Current orientation	0.20	4%

[a] $P = r^2$. Indicates the amount of variation accounted for.

TABLE 7. *Significant social factor effects*

Factor Group	Comments	Factor	Weight	Total Doubled	% Doubled
Language used at work	Much less doubling if all English	Mostly French	0.556	221/447	46
		Mostly English	0.507	155/307	50
		All French	0.409	26/83	31
		All English	0.016	2/21	10
Adult integration	Almost in rank by degree of integration: more integration = more doubling	3	0.910	33/47	70
		1	0.661	143/303	47
		2	0.395	220/472	47
		0	0.161	8/66	12

Two factor groups were of questionable significance, being selected in either step-up or step-down analysis but not in both. The first is sex. Men favored the nonstandard construction more than did women, which parallels a common finding in the L1 studies. However, as we mentioned, the men in this study had lower scores overall for the Acquisition as well as Integration scales, and so this difference is not necessarily attributable to sex. The second is college acquisition. This factor group has the expected pattern. Speakers with higher levels of education in French (those who attended French post-secondary institutions) developed a more standard, less vernacular, French and did not use the doubled subject forms as frequently as those with lower college acquisition scores.

Another question concerns whether more integrated L2 speakers would exhibit a grammar that is like that of L1 speakers or whether they would develop a distinct variety of vernacular French. That is, going beyond the rate of doubling,

TABLE 8. *Comparison of linguistic factors for more and less integrated speakers*

Factor Groups	Factors (in L1 rank order)	High Adult Integration			Low Adult Integration		
		Weight	% Doubled	Total Doubled	Weight	% Doubled	Total Doubled
Clitic type	Ce/Ça >	0.890	85	123/145	0.925	85	81/95
	Personal	0.303	33	120/363	0.290	24	63/267
Subject type	Pronoun >	0.761	68	51/75	0.805	60	28/47
	Proper noun >	0.672	72	47/65	0.552	49	19/39
	Common noun	0.410	39	145/368	0.433	35	97/276
Other preverbal	Adverb >	0.732	55	17/31	0.897	71	12/17
	Hesitation	0.691	67	16/24	0.820	64	9/14
	(Other)	0.681	66	44/67	0.730	46	13/28
	(None)	0.435	43	166/386	0.430	36	110/303
Specificity	Specific >	0.595	57	163/287	0.514	39	106/270
	(Generalizing)	0.387	40	70/175	0.462	46	31/68
	Nonspecific	0.343	22	10/46	0.447	29	7/24
Definiteness	Definite >	0.524	53	219/410	0.540	43	136/315
	Quantified >	0.403	22	17/79	0.275	15	6/39
	Indefinite	0.390	37	7/19	0.171	25	2/8
Clause type	Main >	0.510	49	213/432	0.540	42	131/311
	Subordinate	0.443	39	30/76	0.274	25	13/51
Total		.484	**48%**	**243/508**	.341	**40%**	**144/362**

would more integrated L2 speakers exhibit stronger effects of the linguistic variables that influence L1 speakers? To answer this question, the speakers were split into two groups. Those with a score of 0 or 1 for adult integration (less integrated) were compared to those with a score of 2 or 3 (more integrated). This division was selected because of the significant effect of this social variable in previous analyses.[28] Separate binomial analyses were conducted for higher and lower adult integration speakers. The factor weights from each group and their significance were then compared. The GOLDVARB weights are given in Table 8; factor groups are listed from first to last selected by a GOLDVARB step-up/step-down analysis. Shading indicates the factor groups that were found to be insignificant in each analysis. The factors in each group are listed according to their rank order for L1. Factors not considered in L1 analyses appear in parentheses. Overall doubling rates and input values for the two groups appear at the bottom.

Both groups of speakers exhibited the L1 ranking of factors for each of these factor groups. The exception is other preverbal material, for which neither group illustrated the L1 ranking (as was the case for the analysis of all speakers combined). This analysis reveals the similarity of the grammars of the two groups. Even L2 speakers who have fewer opportunities to speak French in their daily lives had acquired a grammar that largely mirrors that of L2 speakers. Thus, what

TABLE 9. *Doubling rate with subjects that require* ça

	Doubled	Not Doubled	
High doublers	69	6	
Low doublers	19	7	$\chi^2 = 6.16; p \leq .025$

distinguishes highly integrated from less integrated L2 speakers is simply their rate of subject doubling and not their underlying grammar. As the chi-square statistic shows, there is a significant difference in doubling rate between these two groups: $\chi^2 = 5.6$, $p \leq .025$.

It is worth noting that all of the factors that are highly ranked in the L1 analyses are also significant for both groups of speakers. Specificity is selected as significant for speakers with higher indices of integration, whereas definiteness is selected as significant for speakers with lower indices, but this difference does not call the similarity between L1 and L2 into question. Indeed, because there is considerable overlap between the categories of specificity and definiteness, as most specifically interpreted NPs are definite, we may expect GOLDVARB to distinguish poorly between the two factor groups and to select sometimes one, sometimes the other in different analyses.

Contrasting high and low doublers

So far, we have seen that what distinguishes L2 speakers who are well integrated in the Francophone community from those who are less well integrated is their higher rate of subject doubling rather than their underlying grammar. Before concluding, there is one last issue that must be considered. Could it be that this higher rate of subject doubling is not an indication of a more nativelike grammar but rather a reflection of the fact that these speakers use more constructions that favor subject doubling in general? For example, is it possible that highly integrated speakers use more subjects that require doubling with the clitic *ce/ça* (which favors doubling) than do less well integrated speakers, thus resulting in a higher doubling rate? We separated our speakers into two groups (high doublers vs. low doublers) and compared their respective use of constructions that favor subject doubling among L1 speakers; we wanted to see if their different doubling rates could be attributed to their use of such constructions. We found that the different doubling rates could be attributed only in part to the fact that high doublers used more structures that favor doubling.

First, let us consider whether our speakers used equal proportions of subjects that require *ce/ça* as doubling pronouns. Table 9 shows that high doublers had a higher proportion of subjects that require *ce/ça* than low doublers, and that such subjects favor doubling more than those that require *il(s)/elle(s)*. But this is not the whole story. If we compare doubling rates for each type of subject for high and low doublers, we find that the differences are significant. Although

TABLE 10. *Doubling rate with subjects that require* il(s)/elle(s)

	Doubling	No Doubling	
High doublers	71	75	
Low doublers	13	95	$\chi^2 = 37.56; p \leq .001$

TABLE 11. *Doubling rates when* ne *is absent*

	Doubling	No Doubling	
High doublers	139	79	
Low doublers	29	105	$\chi^2 = 59.01; p \leq .001$

high doublers used more subjects that favor doubling, they also doubled these subjects significantly more often than the low doublers did (92% vs. 73%). Additional evidence that high doublers doubled more than low doublers comes from their doubling rates for subjects that require *il(s)/elle(s)* as their doubling clitic (see Table 10). Once again, we see that we have more such subjects for high doublers than for low doublers, but the doubling rate is significantly higher for the former than for the latter (49% vs. 12%).

Similar results were found for other linguistic factors. For negation, the doubling rate when *ne* is absent is significantly different for high and low doublers (see Table 11). It is not when *ne* is present, but this is due to the fact that, for high doublers, there are only three tokens of *ne*. This factor itself illustrates that the French of high doublers was more nativelike: L1 speakers used *ne* less than 2% of the time.

We observed the same scenario for clause type (see Tables 12 and 13). Within each clause type, the difference between high doublers and low doublers is significant. For this reason, the higher rate of high doublers could not be attributed solely to the fact that they used more main clauses than low doublers did.

Finally, although we have few tokens with other preverbal material, the difference between the doubling rates of the two groups is statistically significant for each type of preverbal material. So, once again, the fact that high doublers used more adverbs is not solely responsible for the finding that high doublers doubled more.

One question remains. Is it possible that no single factor results in a significantly higher doubling rate for high doublers than for low doublers, but that it is the accumulation of four different linguistic factors that does? It must be the case that each of these factors plays some role, and that the higher overall percentage for high doublers is due to the fact that they used more subjects that require *ce/ça*,

TABLE 12. *Doubling rates in main clauses*

	Doubling	No Doubling	
High doublers	129	73	
Low doublers	24	91	$\chi^2 = 54.24; p \leq .001$

TABLE 13. *Doubling rates in subordinate clauses*

	Doubling	No Doubling	
High doublers	9	6	
Low doublers	7	48	$\chi^2 = 14.94; p \leq .001$

more adverbs, fewer *ne*, and more main clauses. But these figures suggest that something else is going on—that the sociolinguistic histories of our subjects were in fact affecting their usage of this variable.

There is no significant difference in doubling rate between high doublers and low doublers for subject type, specificity, and definiteness of the subject. These factors do not contribute to the higher rate for high doublers. The lack of a significant difference for these factor groups is compatible with our hypothesis that doubling rates could be attributed to different sociolinguistic histories. It does not, of course, fully confirm it. Given that there are significant differences for four linguistic factors, we cannot rule out the possibility that some linguistic factors contribute to the different doubling rates across speakers.

To summarize, the results presented here confirmed that some L2 speakers had higher rates of subject doubling than others, and that this difference could not be attributed wholly to the fact that high and low doublers used different types of constructions. Combined with the previous results, our findings show that, although all L2 speakers had acquired the main components of subject doubling and knew which factors favor doubling, speakers who were more integrated in the Francophone community better approximated the doubling patterns of L2 speakers by doubling more subjects than their less-well integrated counterparts.

SUMMARY

As a group, these speakers have not completely attained the L1 grammar. Not all linguistic factors were selected as significant, and there were slight discrepancies between the rankings of factors for L1 and L2 speakers. These apparent differences, however, may be due to methodological and geographic differences between this study and those with which it is compared. There are some differ-

ences between how we coded certain linguistic factors and how these were coded in previous L1 studies. Also, we made several comparisons to Nadasdi's work, which examined Ontario French, whereas we dealt with Montreal French. Finally, Nadasdi's study excluded all dislocation examples, whereas others, including ours, did not attempt to distinguish between dislocated and doubled subjects. In spite of these issues, we found much similarity between the L1 and L2 grammars for the variable in question.

A comparison of the behavior of these speakers for other linguistic variables appeared in Blondeau et al. (2002). Briefly, we found very little correlation between speakers' doubling rates and their competence in (virtually) invariant aspects of French that are taught in school: gender marking, verb conjugation, and adverb placement. On the other hand, there was a fair degree of correlation between the rate of subject doubling and the patternings for two phonological variables (variable in L1 as well): the deletion of /l/ in subject pronouns and the affrication of /t/ and /d/ before high vowels.

DISCUSSION

It remains an open question whether these findings indicate partial mastery of the L1 grammar (i.e., a sort of interlanguage status) or whether L2 speakers in this social situation do not wish to acquire L1 patterns that would identify them as Francophone, at the risk of losing their Anglophone identity. We might wish to examine the patterns of Anglophones who identify strongly with Francophone culture and are highly integrated in it, such as Liz and Ted (who both have high overall doubling rates), to test this hypothesis at the extreme. Unfortunately, with only about 40 token phrases available for each speaker, a multivariate analysis of the effects of the linguistic factors on an individual speaker was not possible. To approximate such an approach, we contrasted the less integrated speakers to the more integrated speakers, and we showed quantitative and qualitative differences in their grammars.

As pointed out by Nadasdi (personal communication), the lower doubling rate exhibited by less integrated speakers may well have a structural explanation. For these less advanced speakers, the morphemes we have referred to as clitics may *not* be affixes but rather pronouns occupying an argument position in the syntactic structure, as in English. If so, their lower rate of doubling would be accounted for by the impossibility of putting two elements (the strong pronoun or lexical subject and the "clitic") in the same position. With the pronoun in place, there is no open position for a nonaffixed clitic.

Another possibility is that there is a limit on the number of preverbal affixes that a verb can have in the simpler grammar of the less integrated speakers. Either of these explanations would be supported by the fact that the less integrated speakers rarely used both a subject clitic and any other clitic. For the five speakers with the lowest overall doubling rates, none of the 17 clauses with nonsubject preverbal clitics had doubled subjects. In contrast, for the five speakers with the

highest doubling rate, 10 of 21 (48%) clauses with a nonsubject preverbal clitic had doubled subjects.[29]

Of potential interest to educators and language planners, we found only weak effects of the type of language education on the rate of subject doubling. Neither elementary acquisition nor adolescent acquisition were significant in the binomial regression analysis conducted on the group of speakers as a whole. Of course, there may still be indirect effects. More effective language teaching in schools, including exposure to colloquial Quebec French, may allow students to become more integrated in Francophone culture outside of school (Auger, 2002). This, as we have shown, does correlate with acquisition of this vernacular variable.

This examination of subject doubling in L2 French and the influence of linguistic and social factors shows several things. First, it is clear that all of our L2 speakers had acquired the L1 pattern of subject doubling to some extent, even though it is not explicitly taught in the classroom. Therefore, this is a pattern that is acquired via interaction with L1 speakers. L2 speakers who had more exposure to L1 French as adults did exhibit a higher doubling rate, more closely resembling L1 speakers. All groups of speakers illustrated an effect for most linguistic factors, indicating not only a mastery of a surface pattern, but also an internalization of the grammatically influenced stochastic pattern of L1 speakers.

NOTES

1. Sankoff (1982:85) indicated that doubling rates are very low in formal contexts. She reported that a speaker who doubled 38% of her subjects in an informal context doubled only 5% of them while teaching. We therefore assumed that students do not receive much doubled-subject input in the classroom.

2. Interestingly, the affrication of /t/ and /d/, a feature that displays virtually no variation in L1 French and is associated with local French, is variably adopted by Anglophones. The L2 speakers display a pattern of affrication linked to their degree of contact with a French environment.

3. Although most of the examples given here have singular subjects, our study also included plural subjects.

4. We adopt this abbreviation from Sankoff et al. (1997).

5. In addition to subject doubling, Montreal French permits left dislocation for contrast. The distinction is illustrated by the following pair of sentences (Auger, 1995):

> **Left dislocation:** Les maringouins, ils me suivent '*Mosquitos, they follow me*'.
> **Subject doubling:** Les maringouins ils me suivent '*Mosquitos follow me*'.

The comma in the first example highlights the fact that, for L1 speakers, intonation and liaison may distinguish the two types. In this view, left-dislocated subjects are considered prosodically separate from the following subject and are interpreted as carrying special pragmatic value, such as contrast or emphasis.

6. Note that this doubling rate was calculated differently than in many other studies, including the present study, which only considered sentences containing a lexical or strong pronoun subject and calculated the proportion of subjects accompanied by a clitic subject (e.g., Auger, 1991). The present study reported on the number of subjects consisting of an overt subject (either lexical or pronominal) + subject clitic as a proportion of all subjects: undoubled overt subjects, doubled overt subjects, and clitic subjects.

7. Thanks to Yves Roberge for pointing us toward this finding.

8. For example, clauses whose subject is the strong form of *ça* (which is pronounced with a back vowel, [sɑ], and is clearly distinct from clitic [sa]) were always doubled in this corpus, and so they were not included in the analysis. Sentences containing *sont* 'were.3pl' were considered ambiguous

with respect to doubling, as the overt expression of the pronominal clitic is optional in Quebec French: *(Ils) sont belles ces pommes-là* 'Those apples are beautiful'.

9. Only doubled subjects at the left-edge of the sentence were counted. Right-dislocations were coded as single subjects, since they could not possibly be instances of the double marking phenomenon we were examining.

10. Nadasdi (1995a), for example, excluded all instances of left-dislocation and all cases for which a dislocated interpretation could not absolutely be ruled out. In other words, only clear cases of subject doubling were analyzed.

11. This detailed version was included to compare how L2 subjects react to the degree of specificity of the subject, as examined by Nadasdi (1995a) for L1 speakers. However, we added the variants "verb" and "impersonal pronoun," which Nadasdi did not use, and we eliminated his "indefinite noun" factor in favor of a separate factor group for definiteness.

12. Nadasdi's (1995a) study did not consider phrases with *ça* as part of the variable context.

13. This variant was not included in Nadasdi (1995a).

14. Nadasdi (1995a) examined the effect of the preverbal negative marker *ne* as part of the factor group preverbal clitics. As we found a significant effect for *ne* but not for other preverbal clitics, we decided to examine them separately.

15. Nadasdi (1995a) reported effects of hesitations and adverbs but did not examine the other preverbal elements that we included.

16. Nadasdi (1995a) separated the expression *tout le monde* 'everyone', which had a low doubling rate (5% doubled) compared to other quantified expressions (18% doubled) in his sample. We did not make this distinction. 45 of 148 tokens we coded as "quantified" consisted of *tout le monde*.

17. We added this variant because of the possibility that doubling might be more frequent with generalizing NPs. Indeed, as *ce/ça* often forces a generic interpretation and removes ambiguity, we can imagine that this type of NP favors doubling. Compare the following sentences:

(i) *Les jeunes d'aujourd'hui fument plus*.
(ii) *Les jeunes d'aujourd'hui ça fument plus*.
     'Young people no longer smoke'

However, the L1 data to which we compared our L2 data did not include this variant.

18. That is, female speakers in our sample (and, we believe, in the population) had higher average scores for both acquisition and integration. But, although women exhibited a higher doubling rate, it could not be attributed directly to their being female. It would be equally valid to attribute the higher doubling rate to greater integration or higher levels of acquisition. Thus, we did not include sex as a social factor in our analyses.

19. Because of the homogeneity of our sample, we could not adequately measure the effects of social class or linguistic marketplace. Nearly all of the speakers came from middle class families, and many were still students and so are not yet engaged in work that defines their identity.

20. Fractional scores indicate partial attendance in French or immersion programs. For example, a score of 1.5 reflects partial attendance of one to two years in a French or immersion elementary school; 2 points are attributed for three to four years; and 2.5 points are given for attendance in post-immersion programs starting after the fifth grade.

21. In some cases, half points were assigned if the speaker had done a small amount of an activity that involved integration with French culture.

22. These doubling rates are based on our full corpus of 1,100 tokens.

23. This ranking is based on GOLDVARB's step-up/step-down analysis.

24. The Anglophones exhibited a 90% rate of *ne* absence (Blondeau et al., 2002). This approaches the 99% rate observed for spoken L1 French in Montreal (Sankoff & Vincent, 1977).

25. Nadasdi also examined the effect of verb type in Ontario French using a similar coding system to ours: transitive, unergative, unaccusative, and passive (but no *être*). Nadasdi (1995a:113) found, as we did, that transitive and unergative verbs favor doubling. Nadasdi did not expect to find a significant difference between passives and unaccusatives, but he did. He had expected a pattern such as the one found in our study. Although this group is not significant in our data, the ranking is similar to what Nadasdi found for L1 speakers. Nadasdi showed an effect of preverbal clitic in Ontario French. However, his preverbal clitic factor group included *ne*, object clitics, and reflexives. We considered *ne* as a separate factor (preverbal negation) in our analysis and found it to have a significant effect.

26. We also included number as a factor group in this analysis, but it was not significant.

27. For the purposes of creating this graph, the doubling rate was divided by 20 to make it fit on the same scale as the other variables. Overall, 1,100 tokens were considered; thus these percentages

differ slightly from those shown in Table 5, which were calculated from a smaller file that did not contain knock-out factors.

28. Other comparisons of this type, such as contrasting speakers who work in French versus English, were not possible due to small numbers of speakers in certain categories.

29. This difference is significant, $\chi^2 = 11$, $p < .001$.

REFERENCES

Auger, Julie. (1991). *Variation and syntactic theory: Agreement-marking vs. dislocation in Quebec Colloquial French*. Paper presented at NWAVE-20, Washington, DC.

———— (1994). *Pronominal clitics in colloquial Québec French: A morphological approach*. Doctoral dissertation, University of Pennsylvania. (Distributed by the Institute for Research in Cognitive Science, University of Pennsylvania, IRCS Report No. 94-29.)

———— (1995). A morphological analysis of Quebec Colloquial French pronominal clitics. In A. Dainora et al. (eds.), *CLS 31-II: Papers from the parasession on clitics*. Chicago: Chicago Linguistic Society. 32– 49.

———— (1998). Le redoublement des sujets en français informel québécois: Une approche variationiste. *Canadian Journal of Linguistics* 43:37– 63.

———— (2002). French immersion in Montreal: Pedagogical norm and functional competence. In Kathleen Bardovi-Harlig, Susan Gass, Sally Magnan, & Joel Walz (eds.), *Pedagogical norms for second and foreign language learning and teaching*. Amsterdam: Benjamins. 81–101.

Beaulieu, Louise, & Balcom, Patricia. (1998). Le statut des pronoms personnels sujets en français acadien du nord-est du Nouveau-Brunswick. *Linguistica Atlantica* 20:1–27.

Blondeau, Hélène. (1999). *Parcours d'un marqueur sociolinguistique. Les pronoms non-clitiques en français parlé de Montreal*. Doctoral dissertation, Université de Montréal.

Blondeau, Hélène, Nagy, Naomi, Sankoff, Gillian, & Thibault, Pierrette. (2002). La couleur locale du français L2 des anglo-Montréalais. In J.-M. Dewaele & R. Mougeon (eds.), *L'acquisition de la variation par les apprenants avancés du français langue seconde: Acquisition et Interaction en Langue Étrangère* 17:73–100.

Caroll, Susan. (1982). Redoublement et dislocations en français populaire. In C. Lefebvre (ed.), *La syntaxe comparée du français standard et populaire: Approches formelles et fonctionnelles* (Vol. 1). Québec: Gouvernement du Québec. Office de la Langue Française. 290–357.

Deshaies, Denise, Guilbault, Christian, & Paradis, Claude. (1992). Prosodie et dislocation à gauche par anaphore en français québécois spontané. In A. Crochetière, J. C. Boulanger & C. Ouellon (eds.), *Actes du XVième Congrès international de linguistique*. 31–34.

Genesee, F. (1987). *Learning through two languages: Studies of immersion and bilingual education*. Cambridge: Newbury House.

Givón, Talmy. (1976). Topic, pronoun, and agreement. In C. Li (ed.), *Subject and topic*. New York: Academic. 148–186.

Grevisse, Maurice, & Goosse, André (ed.), (1986). *Le bon usage: Grammaire française*. Paris: Duculot.

Kayne, Richard. (1977). *Syntaxe du français: Le cycle transformationnel*. Paris: Seuil.

Nadasdi, Terry. (1995a). *Variation morphosyntaxique et langue minoritaire: Le cas du français ontarien*. Doctoral dissertation, University of Toronto.

———— (1995b). Subject NP doubling, matching, and minority French. *Language Variation and Change* 7:1–14.

Nagy, Naomi, & Blondeau, Hélène. (1998). Double subject marking in L2 Montreal French. *University of Pennsylvania Working Papers in Linguistics* 6:93–108.

Nagy, Naomi, Moisset, Christine, & Sankoff, Gillian. (1996). On the acquisition of variable phonology in L2. *University of Pennsylvania Working Papers in Linguistic* 3:11–126.

Rand, David, & Sankoff, David. (1990). GOLDVARB version 2: A variable rule application for the Macintosh. Montréal: Centre de recherches mathématiques, Université de Montréal.

Rehner, Katherine, & Mougeon, Raymond. (1999). Variation in the spoken French of immersion students: To *ne* or not to *ne*, that is the sociolinguistic question. *The Canadian Modern Language Review/La Revue Canadienne des Langues Vivantes* 56:124–154.

Rehner, Katherine, Mougeon, Raymond, & Nadasdi, Terry. (1999). *Variation in the spoken French of immersion students:* Nous *versus* on. Paper presented at the Second Language Research Forum, University of Minnesota, Minneapolis.

Sankoff, Gillian. (1982). Usage linguistique et grammaticalisation: Les clitiques sujets en français. In N. Dittmar & B. Schlieben-Lange (eds.), *La sociolinguistique dans les pays de langue romane*. Tübingen: Gunter Narr Verlag. 81–85.

Sankoff, Gillian, Thibault, Pierrette, Nagy, Naomi, Blondeau, Hélène, Fonollosa, Marie-Odile, & Gagnon, Lucie. (1997). Variation in the use of discourse markers in a language contact situation. *Language Variation and Change* 9:191–218.

Sankoff, Gillian, & Vincent, Diane. (1977). L'emploi productif du *ne* dans le français de Montreal. *Le Français Moderne* 45:243–256.

Southard, Bruce, & Muller, Al. (1998). Blame it on Twain: Reading American dialects in *The Adventures of Huckleberry Finn*. In Dalin Oaks (ed.), *Linguistics at work: A reader of applications*. Harcourt Brace. 565–573.

Thibault, Pierrette. (1983). *Équivalence et grammaticalisation*. Doctoral dissertation, Université de Montréal.

Thibault, Pierrette, & Sankoff, Gillian. (1997). *The insertion of a second language into the community repertoire: Anglophone French in Montreal*. Paper presented at NWAVE-26, Université Laval.

———— (1999). L'évaluation du français des jeunes Anglo-Montréalais par leurs pairs francophones. *The Canadian Modern Language Review/La Revue Canadienne des Langues Vivantes* 56:245–281.

———— (1993). Diverses facettes de l'insécurité linguistique. *Cahier de l'Institut de linguistique de Louvain* 19:209–218.

Wolfram, Walt, & Christian, Donna. (1976). *Appalachian speech*. Arlington, VA: Center for Applied Linguistics.