Subject NP doubling, matching, and minority French

TERRY NADASDI
York University

ABSTRACT

Our study presents a variationist analysis of subject doubling in the French of Ontario, Canada. Two principal variants are distinguished: a non-doubled variant and a doubled variant containing a clitic agreement marker. In our analyses, both linguistic and social factors are taken into account and analyzed using GOLDVARB2. It is proposed that subject clitics are marked for default features, and that the doubled variant is favored when the clitic's default features match those of the subject NP; lack of matching favors the non-doubled variant. Discussion of linguistic factors for the present study, therefore, is limited to those factors which can be explained in terms of matching. The principal social factor studied is restricted language use (cf. Mougeon & Beniak, 1991). Our results show that the greater the restriction, the fewer doubled subjects one finds.

Subject doubling constructions (hereafter SD) in French have recently received a great deal of attention within the various subparadigms of generative grammar and variationist studies (cf. Auger, 1991; Carroll, 1982; Ossipov, 1990; Roberge, 1990; Roberge & Vinet, 1989; Sankoff, 1982). Still, few have attempted to verify their findings on a large body of spoken speech or to take into account both linguistic and social factors. Furthermore, there has been no attempt to examine the correlations between subject clitic doubling and minority language restriction. The present study not only concentrates on this latter domain, but also serves to demonstrate the need for, and advantages of, quantitative approaches to linguistic constraints on doubling phenomena. We show, for example, that contextual factors involving the matching of features like [±specific] and [±subject] between a clitic and a subject NP are often quantitative rather than categorical. The relevance of such features has been suggested in previous work on formal syntax, but their variable effect on SD has remained unexplored and, as such, poorly understood.

CORPUS

The following analyses are based on the spoken French of 117 adolescent speakers residing in Eastern and Central Ontario. Four communities are rep-
resented: Pembroke, North Bay, Cornwall, and Hawkesbury. All speakers have at least one francophone parent and have been schooled in French. But an important difference separates the speakers of Hawkesbury from those in the other three communities: in Pembroke, North Bay, and Cornwall, francophones constitute a linguistic minority, the majority of speakers in these areas being anglophone. The present study will examine the impact of this on the French spoken in these localities.

IDENTIFYING THE VARIABLE

The morphosyntactic variable under study can be generally stated as the presence or absence of SD: that is, the presence or absence of a subject clitic within a clause containing a subject NP. In the general category of "subject NP," we include referential subjects, such as la table, Jean, un ami à moi, as well as strong pronominal subjects, such as moi, toi, lui, eux-autres, etc. The motivation for this classification is that in both cases we are dealing with an independent NP functioning as subject. Subject NP is then used in contrast with what could be termed "subject affix," a class which includes pronominal subject clitics, such as je, tu, il, etc. Examples of each variant are presented in (1) and (2):

(1) a. Steve il tombe en amour avec. (Cor-05)
b. Mon père _____ m’emmène au Mont Tremblant. (Haw-19)
(2) a. La boucane elle sortait du plancher. (Cor-26)
b. Les écoles _____ vont être plus sévères. (Cor-24)

The (b) variant, without the clitic, corresponds to what one normally finds in Standard French; the (a) variant, which is an instance of SD, is found in various nonstandard varieties (cf. Auger, 1991; Carroll, 1982; Roberge, 1990), including the variety of French spoken in Ontario (hereafter FO).

In the quantitative analyses that follow, only third-person forms are taken into account since clauses containing a first- or second-person strong form must always include a subject clitic, as shown in (3):

(3) a. *Moi travaille demain.
b. Moi je travaille demain.
c. *Toi veux partir?
d. Toi tu veux partir?

It should be pointed out that the reverse constraint does not hold: that is, the presence of a subject clitic does not require that there be a strong form in the same clause. As such, it is possible to approach SD from a different perspective by examining the presence or absence of a subject NP in clauses contain-
ing a subject clitic. However, this would constitute a different variable than the one studied here, as it gives rise to different kinds of alternations.

Our reason for examining doubling in terms of the presence or absence of a subject clitic is that this approach would allow us to examine several hypotheses concerning the correlations between the presence or absence of bound morphemes and language restriction. The principle hypothesis we wanted to test was first put forth by Andersen (1982:102):

An LA [i.e., a restricted speaker] will show a marked preference for syntactic constructions using free morphemes where there is a choice between free and bound morphemes in comparison with an LC [i.e., an unrestricted speaker].

Since the doubled variant contains a bound morpheme (i.e., the pronominal clitic), we hypothesized that this structure would be used less frequently by restricted speakers. While the non-doubled variant does not contain a free form other than the subject NP, the fact that it provides the speaker with a means of not using a bound morpheme should result in more frequent use of this structure as language restriction increases.

SUBJECT DOUBLING VERSUS LEFT-DISLOCATION

Let us first consider the structural properties of SD. SD resembles a structure often referred to as left-dislocation. In strictly linear terms, both involve the presence of a subject NP and a subject clitic to the left of the verb (cf. Ashby, 1988; Barnes, 1985, 1986; Hirschbühl, 1975; Lambrecht, 1981); however, the two can be distinguished by several tests, as exemplified in Roberge (1990). On the prosodic level, these two kinds of “double” subjects are distinguishable in that a left-dislocated subject may be separated from the verb by a pause and can bear contrastive or emphatic stress. Furthermore, unlike what one finds with SD, there is never linking between the final consonant of the left-dislocated NP and a following vowel (cf. Deshaies, Guibault, & Paradis, 1992). As such, it can be structurally analyzed as occupying a topic position, whereas the subject in a doubling construction is found in the structural position of subject. The structural representation of SD is given in (4), and that of left-dislocation is presented in (5): 4

(4)

```
C'  
\    \  
C    IP  
\    \  
NP    I'  
\    \  
mon fils I    VP
  \  
    il travailler
```
These two structures—taken from Roberge (1990) and based on Rizzi (1986) and Chomsky (1977)—show that the NP of a left-dislocation occupies the position TOP rather than that of structural subject (i.e., NP of IP). The validity of this analysis is confirmed by the fact that, when there are elements in COMP (e.g., wh-words), the NP of a left-dislocation is always to the left of these elements:

(6) a. Ton frère là, quand qu’il va rentrer?
    b. Il a fallu, maman, qu’elle aille travailler sur la ferme.
    c. *Ton frère quand qu’il va rentrer.
    d. *Il a fallu maman qu’elle aille travailler sur la ferme.

We can conclude that, structurally, the NP of a left-dislocation does not occupy the syntactic position of verbal subject.

In order to distinguish instances of SD from cases of left-dislocation, all potential occurrences of the variable were listened to by the author to determine whether the prosodic features (e.g., pause, emphatic or contrastive stress, absence of consonantal linking) were present. If so, these occurrences were eliminated from the occurrences of the variable since, as we have just seen, left-dislocated subjects are clearly not structural equivalents of SD.

For the present study, we looked at almost 3,000 tokens, coding various linguistic factors based on our own observations, as well as on previous work on SD by Laurendeau, Fournier, and Néron (1981), Sankoff (1982), Carroll (1982), Roberge (1990), and Auger (1991). Our purpose was to determine the relative contribution, if any, of various social and linguistic factors in SD structures.

LINGUISTIC FACTORS

Specificity and definiteness

The first factor we examined was the specificity and/or definiteness of the subject NP. Depending on the analysis (cf. Chafe, 1976; Chesterman, 1991;
Christophersen, 1939; Enc, 1986; Hawkins, 1978; Heim, 1982; Jespersen, 1949; Rando & Napoli, 1978; Reuland & Meulen, 1987), specificity and definiteness can be viewed as either distinct from one another or as overlapping. Our initial approach was to keep them separate, using two binary features: [±specific] and [±definite]. Adopting a straightforward test found in most of the work on definiteness, we considered any NP that occurred in an existential or presentative sentence to be [−definite]. For practical purposes, this meant that it was able to function as the object of il y a. The class of [−definite] nouns included all nouns that were preceded by an indefinite article, as well as the traditional indefinite pronouns (e.g., quelqu’un, autre chose, plusieurs, certains). By the il y a test, [+definite] nouns were those preceded by a definite article, as well as proper nouns and personal pronouns.

Regarding specificity, we adopted the definition used by Suñer (1988): [+specific] NPs were those whose referent could be identified. Examples of [±definite] and [±specific] NPs are given in examples (7) through (10):

(7) +definite
   a. Non, i’ disent que les gars ils aiment ça des filles courtes. (Nor-11)
      No, they say that guys they like that short girls.
   b. Je pense que les jeunes veulent trop. (Nor-08)
      I think that young people want too much.

(8) −definite
   a. Un homme il dit à l’homme . . . (Pem-15)
      A man he said to the man . . .
   b. Plusieurs Anglais apprennent le français à l’école. (Cor-10)
      Several English people learn French in school.

(9) +specific
   a. Mes parents ils étaient partis. (Nor-05)
      My parents they left.
   b. J’ai monté souvent quand mon père travaillait sur les affaires pour le feu là. (Pem-35)
      I often went there when my father was working on things for fires there.

(10) −specific
   a. En Ontario les gens ils hésitent. (Nor-19)
      In Ontario people they hesitate.
   b. Souvent pour moi c’est parce que les enfants parlent pas assez avec leurs parents. (Pem-35)
      In my opinion, it’s often because children don’t talk enough with their parents.

Our reason for including the feature [±definite] is based on Roberge’s (1990) claim that, in Colloquial French, SD is impossible with indefinite NPs. He cited the sentence in (11)

(11) *Un homme il vient. [Roberge, 1990:63]
as an example of ungrammatical doubling due to indefiniteness. Now, this is clearly not the case for FO, since indefinites can certainly be doubled in this variety, as illustrated in (12):

(12) a. *Un Francais de France* il va arriver au Canada . . . (Cor-10)
    b. *Un groupe de personnes* il essaie de tuer un inspecteur. (Cor-10)
    c. *Des personnes* ils boit. (Pem-34)

Still, it seems plausible that, if the constraint on indefinites is in fact categorical in Colloquial French, it may also have a quantitative effect on doubling in FO.

The [±specific] distinction was included since, given the sentences in (12), it may be that [−definite] does not inhibit doubling, but that [−specific] does. This seems likely if personal pronouns are normally [+specific] (cf. Suñer, 1988; Thibault, 1983).

Our first step then was to code all tokens for these binary features. Our initial approach was that, if plus values for each group contributed in a significant way, they would subsequently be combined, much as Quirk, Greenbaum, Leech, and Svartvik (1972) did for their grammar of English. As both specificity and definiteness were selected when examined individually, they were then combined. The GOLDVARB results for these combinations are presented in Table 1. As can be seen by the factor weights given here, SD is more likely to occur when the subject NP is [−specific] and [−definite]. It is clearly inhibited when the NP has negative values for these features.

Why though should this be the case? We would suggest that this can be explained by appealing to Suñer’s (1988) matching hypothesis, which she used to account for object doubling in Spanish. Suñer claimed that the occurrence of a doubling object clitic depends on the matching of such features as [animate] and [specific] between the clitic and the NP sharing the same syntactic function. The Matching Hypothesis states that

CLs [clitics] are inflections, IO [indirect object] and DO [direct object] CLs are defined by assigning values to the features [specific], [animate] . . . Once a CL, having a referential index, becomes part of the verbal head, it must match not only the referential index transferred from one of the complements to the theta-grid of the V, but also the relevant features of the complement. A chain is well formed only when there is no clash in features. (Suñer, 1988:402–403)

As for matching and SD, we hypothesized that the occurrence of a subject clitic within the same clause as a subject NP (i.e., subject doubling) would depend on the matching of the personal pronoun clitic’s [+specific] value with that of the subject NP. We considered personal clitics [+specific] by default, in contrast with the subject clitic *ça*, which is normally [−specific]. This view is supported by Thibault (1983), who showed that in Montréal French [+spe-
TABLE 1. Subject doubling and specificity

<table>
<thead>
<tr>
<th></th>
<th>Factor Weight</th>
<th>Tokens</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>+spec./+def.</td>
<td>.587</td>
<td>529/1513</td>
<td>35</td>
</tr>
<tr>
<td>+spec./−def.</td>
<td>.575</td>
<td>13/44</td>
<td>30</td>
</tr>
<tr>
<td>−spec./+def.</td>
<td>.408</td>
<td>150/920</td>
<td>16</td>
</tr>
<tr>
<td>−spec./−def.</td>
<td>.186</td>
<td>14/138</td>
<td>10</td>
</tr>
</tbody>
</table>

Specific subjects are usually represented by a personal subject pronoun, whereas those having a negative value for this feature are almost always represented by ça/c'.

Suñer proposed that matching applies categorically for Spanish object doubling. However, according to our data, the matching hypothesis does not always obtain for SD in FO. Still, the data in Table 1 suggest that it is valid as a quantitative constraint: that is, the more this constraint is violated, the less likely one is to find a doubled subject. In other words, the likelihood of finding a [+specific] marker (i.e., the personal pronominal clitic) greatly diminishes when the coreferential NP is [−specific]. Further research should show that doubling by a [−specific] marker would be favored by NPs bearing this same value (cf. Auger, 1993).

It would be possible to leave the analysis of specificity here with specificity as a binary feature, its presence being more favorable to doubling than its absence. However, a number of studies on specificity, and on definiteness when this encompasses specificity, suggest that this feature should be viewed as scalar. Comrie (1981), for example, posited a continuum of specificity; Chesterman (1991) pointed out that there are in fact various degrees of definiteness. This same idea is also present in Quirk et al. (1972), who distinguished [+specific] common nouns from [+specific] proper nouns.

Given this, it was hypothesized that, since SD is more frequent with [+specific] NPs, it would be found most often with NPs occupying a marked position on a continuum of specificity. The continuum used to verify this is shown in (13):

(13) Specificity continuum

<table>
<thead>
<tr>
<th>1st and 2nd person</th>
<th>3rd person</th>
<th>proper nouns</th>
<th>common nouns</th>
<th>indefinite pronouns</th>
</tr>
</thead>
</table>

The results of the GOLDBARB analysis for this factor group are given in Table 2. As mentioned before, first- and second-person strong subject pronouns were not included for coding since they always give rise to the doubled variant in nonstandard varieties. However, they have been included here as they rep-
TABLE 2. Subject doubling according to degree of specificity of the subject NP

<table>
<thead>
<tr>
<th>Noun Type</th>
<th>Factor Weight</th>
<th>Tokens</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st &amp; 2nd person pronoun</td>
<td>KO</td>
<td>—</td>
<td>100</td>
</tr>
<tr>
<td>3rd person pronoun</td>
<td>.861</td>
<td>145/195</td>
<td>74</td>
</tr>
<tr>
<td>Proper noun</td>
<td>.681</td>
<td>51/118</td>
<td>43</td>
</tr>
<tr>
<td>Common noun</td>
<td>.462</td>
<td>496/2187</td>
<td>23</td>
</tr>
<tr>
<td>Indefinite pronoun</td>
<td>.281</td>
<td>14/115</td>
<td>12</td>
</tr>
</tbody>
</table>

Note: Significance = .009.

resent the extreme left of a specificity continuum. The results in Table 2 are in line with our hypothesis regarding doubling and degree of specificity: the further one moves to the right, the less likely one is to find a doubled subject. As well as supporting the view that specificity should be viewed as scalar rather than binary, the information in Table 2 offers an explanation for why first- and second-person strong pronouns never occur undoubled in non-standard French. Their high degree of specificity, attributable to the speaker's or listener's presence at the act of communication, forces the use of a personal subject clitic. In other words, matching is absolute.

Initial position of the subject NP

The second factor group we examined was the underlying position of the subject. Three possibilities were distinguished: (a) initial subject, (b) initial object of an unaccusative verb, and (c) initial object in a passive construction.

In the case of initial subjects, we included all agentive and non-agentive transitives, as well as subjects of intransitive unergative verbs. Following Perlmutter (1978), we considered unergative those intransitive verbs whose subject was in some way responsible for the action of the verb. Two categories of unergatives were distinguished: (a) predicates describing volitional acts (travailler, jouer, parler, sourire, etc.), and (b) predicates describing certain corporal processes (pleurer, éternuer, dormir, tousser, etc.). Examples of the variable with transitives and unergatives are given in (14) and (15):

(14) transitive
   a. Le gars il a lâché l'école. (Haw-20)
   b. Mon père ______ m'emmène au Mont Tremblant. (Haw-19)

(15) unergative
   a. la petite fille elle pleurait. (Nor-19)
   b. ... dans une cave où mon frère ______ dormait. (Pem-15)
In the class of initial object, we placed the arguments of passives, as well as what Perlmutter (1978) analyzed as unaccusatives. In the latter class, one finds intransitive verbs whose subject does not actively participate in the action described. Traditional unaccusative verbs then are those whose subject functions semantically as a patient (e.g., brûler, tomber, glisser), verbs of existence (e.g., exister, se passer), and aspectual predicates (e.g., commencer, arrêter, cesser). Examples of the variable with unaccusatives and passives are given in (16) and (17):

(16) unaccusative
   a. Le gars il tombe en bas. (Cor-26)
   b. On avait une vieille tente que la pluie rentrait dedans. (Haw-12)
(17) passive
   a. Eux-autres ils ont été forcés un peu d’aller à la messe. (Nor-18)
   b. Mon père a été élevé sur une ferme. (Nor-18)

Our reason for examining the influence of the initial position of subject on SD was quite simple. We hypothesized that, since subjects of passives and unaccusatives resemble objects, they would be less likely to give rise to the doubled variant, as the doubled variant contains a subject marker—the subject clitic. The [+subject] feature borne by these clitics can be formalized along the lines of Cummins and Roberge (1994), who proposed that morphosyntactic features are attributed to verbal morphemes in a morphological component of the grammar. With regard to SD, we could say that the occurrence of a clitic bearing a [+subject] feature with an initial object would constitute a semantic or logical mismatch.

The results for this factor group are given in Table 3. In our first GOLDVARB2 run, both non-agentive transitives and unergative subjects were separated. However, since the factor weights favoring doubling were quite similar to those of agentive transitives, the three types were collapsed. This is justified since they are structural equivalents: they are all underlying or initial subjects. On the other hand, the difference between unaccusatives and passives was selected as significant.

The results from Table 3 show that both unaccusative and passive constructions inhibit doubling. We would suggest that it is possible to explain this in terms of a mismatch between the subject clitic and the derived subject NP of unaccusatives and passives. The use of a subject clitic within the same clause as an object-initial NP represents a kind of mismatch between the [+subject] feature of the clitic and the “object-like” NP of an unaccusative or passive—not unlike that between a [+specific] personal clitic and a [−specific] NP. Although the NPs of passives and unaccusatives surface as syntactic subjects, the object status of these patients continues to make its presence felt in quantitative terms. We would posit further that this mismatch is most glaring with passives, which would explain why this group least favors SD.
### Table 3. Subject doubling and initial position of subject

<table>
<thead>
<tr>
<th>Factor</th>
<th>Weight</th>
<th>Tokens</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject transitive/unergative</td>
<td>.607</td>
<td>477/1306</td>
<td>37</td>
</tr>
<tr>
<td>Unaccusative subject</td>
<td>.414</td>
<td>215/1160</td>
<td>19</td>
</tr>
<tr>
<td>Passive subject</td>
<td>.251</td>
<td>14/149</td>
<td>9</td>
</tr>
</tbody>
</table>

### SOCIAL FACTORS

#### Language restriction

One of the central goals of this study was to determine if variables containing pronominal clitics are influenced by restriction in language use. In the corpus on which our analyses were based, speakers were given an index of language restriction based on their language use habits, as revealed through information obtained by questionnaires. Three levels of restriction were distinguished: highly restricted (.05-.44), semi-restricted (.45-.78), and unrestricted (.8-1). It was possible to consider these indices as levels of French language maintenance: the lower the index, the less often the speaker uses French. These indices are explained in detail in Mougeon and Beniak (1991).

If we now examine Table 4, we can see that language restriction correlates well with SD: the greater the restriction (i.e., the less often a speaker uses French), the fewer doubled subjects. The question then is, why should this be so? Several explanations seem possible.

**Linguistic convergence.** The first explanation appeals to the potential influence of English on the French of restricted speakers. The basic claim is that—since restricted speakers speak English fluently and use it often, and SD (or weak subject pronouns, for that matter) is not found in English—restricted speakers avoid doubled subjects when speaking French. This then would be a case of covert interference: a variant not existing in the stronger language is under-used when speaking the weaker language (cf. Mougeon & Beniak, 1991).

However, an external explanation for the relatively high frequency of the undoubled variant among restricted speakers quickly reaches an impasse. It is possible to analyze this variant as simpler, given that it contains fewer elements than the doubled one. This is perhaps contentious (cf. Martinet, 1969:195), but if it is true the difficulties of proving an influence claim are insurmountable (cf. Cassano, 1977; Nadasdi, 1989); the general tendency of restricted speakers is to use simpler or more regular structures, regardless of the dominant language.

**Stylistic reduction.** Another possible explanation linked to language restriction is the reduction in stylistic variation, a feature which often char-
TABLE 4. Subject doubling and language restriction

<table>
<thead>
<tr>
<th>Degree of Restriction</th>
<th>Factor Weight</th>
<th>Tokens</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>.80-1</td>
<td>.662</td>
<td>298/733</td>
<td>41</td>
</tr>
<tr>
<td>.45-.78</td>
<td>.499</td>
<td>276/1116</td>
<td>25</td>
</tr>
<tr>
<td>.05-.44</td>
<td>.346</td>
<td>132/766</td>
<td>17</td>
</tr>
</tbody>
</table>

TABLE 5. Subject doubling and social class in Hawkesbury

<table>
<thead>
<tr>
<th>Social Class</th>
<th>Factor Weight</th>
<th>Tokens</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Working</td>
<td>.698</td>
<td>61/114</td>
<td>54</td>
</tr>
<tr>
<td>Lower middle</td>
<td>.430</td>
<td>48/158</td>
<td>30</td>
</tr>
<tr>
<td>Middle</td>
<td>.339</td>
<td>19/76</td>
<td>25</td>
</tr>
</tbody>
</table>

characterizes minority languages. As Andersen (1982:112) put it, “a restricted speaker will exhibit a narrower range of styles, registers, ways of speaking, etc., than a comparable unrestricted speaker.” Among adolescent Franco-Ontarians schooled in French, stylistic reduction takes the form of the loss of nonstandard variants; the use of French by many of these speakers is limited to a scholastic setting, where Standard French is the observed model. This view is in fact supported by the data. Only in the majority francophone community of Hawkesbury was there a correlation between social class and SD. The relevant figures for SD according to social class in Hawkesbury are presented in Table 5. Since SD is characteristic of Colloquial French, its absence among restricted speakers could be explained by the fact that these speakers had less exposure to the vernacular and thus did not have full command of this register. It is not surprising then that social class does not condition the variable in the minority communities.

Bound morpheme avoidance/morphological simplification. Although stylistic reduction does provide a partial explanation for diminished use of SD among the restricted speakers, it is incomplete, since other clitic variants that are not particular to the vernacular are also preferred by restricted speakers. For example, Nadasdi (1995) showed that this same pattern obtains for object clitics where the use of a strong form cannot be explained as a stylistic reduction since it characterizes neither the standard nor the vernacular variety. For example,

(18) Quand je leur parle là, i’ comprend moi en français. (Pem-16)
Let us therefore consider a third explanation, which stems from Andersen's (1982) hypothesis: that is, restricted speakers will tend to prefer free morphemes compared to bound ones. There is, however, a noticeable difference between the situation Andersen described and the absence of SD among restricted speakers of FO. In the case of SD we are not dealing with a choice between a free morpheme and a bound morpheme, but rather between a bound morpheme and nothing. Thus, we would modify Andersen's hypothesis by stating that restricted speakers show a marked preference for structures which allow them to avoid bound morphemes where possible.

It should be observed that morphological simplification among restricted speakers has also been noted by Mougeon and Beniak (1991) in their study of third-person plural verbal morphology. These authors argued that restricted speakers demonstrate a strong tendency to merge third-person plural agreement with third-person singular forms. Avoidance of SD is parallel to this; in both cases, a verbal morpheme frequently used by unrestricted speakers is used less often by restricted speakers. It is also worth noting that the parallel between avoidance of subject clitics and loss of third-person plural morphology provides external evidence to support the claim that subject clitics are best analyzed as verbal affixes.

CONCLUSION

We suggest that SD in FO is conditioned by linguistic factors, which can be explained by giving a quantitative slant to Suñer's matching hypothesis. Our results indicate that claims regarding categoric constraints on syntactic structures need to be revisited and considered in quantitative terms, using data from a corpus of spoken language. Regarding language restriction, we have shown that the higher the level of restriction, the fewer doubled subjects one finds. This can be linked to three possible sources: (a) external influence, (b) stylistic reduction, and (c) bound morpheme avoidance. However, only the third possibility offers a complete explanation of the phenomenon in question.

NOTES

1. These speakers represent part of a larger corpus established by Raymond Mougeon and his team around 1980. For details concerning the methodology used in gathering these data, see Mougeon and Beniak (1991).
2. Codes refer to locality and speaker number (Cor = Cornwall, Haw = Hawkesbury, Nor = North Bay, Pern = Pembroke).
3. While non-doubled sequences involving first-person forms are possible in the standard variety (e.g., "Ma soeur et moi partons demain"), this structure is not found in Ontario French and, to the best of our knowledge, is also absent from other nonstandard varieties.
4. For an overview of the structural representation of topic NPs, see Rochement (1989).
5. This example is from Roberge (1990:94).
6. Given the specific goals of this article, our discussion of linguistic factors influencing the variable will be limited to those that can be interpreted in terms of “matching.” For a detailed discussion of other relevant factor groups, see Nadasdi (1995).
7. Note that Roberge’s claim regarding absence of doubling with indefinite NPs is not substantiated by corpus studies and is therefore subject to empirical confirmation.
8. It should be pointed out that the specificity and definiteness factor groups were analyzed in separate GOLDVARB runs, given the potential interaction between these groups.
9. As discussed in Nadasdi (1995:121), clitic doubling is also more frequent with [+animate] NPs.
10. As an anonymous review has suggested, the differences between unaccusatives and passives could be related to the fact that speakers may be better able to identify an object in the case of passives, and that this could be linked to their ability to associate passives with their active counterparts, an option not available with unaccusatives.

REFERENCES


