Variation in subject doubling in Homeland and Heritage Faetar

Katharina Pabst¹, Lex Konnelly¹, Fiona Wilson¹, Savannah Meslin², and Naomi Nagy¹

¹University of Toronto, ²Canada’s National Ballet School

This paper investigates subject doubling in Faetar, an endangered and understudied variety of Francoprovençal. Comparing Homeland speakers (i.e., speakers who were born and raised in Faeto) and Heritage speakers of the language (i.e., speakers who emigrated to Toronto, Canada after age 18, and their children), we find some striking differences. Our results show that subject doubling is grammatically constrained in the source variety: Homeland speakers favor doubling in new information contexts, while Heritage speakers do not. There is also evidence for a change in progress among Homeland speakers, with younger speakers using more subject doubling than older speakers. This change is not mirrored by the Heritage speakers. We propose that this is because the Heritage speakers left the Homeland either before or around the time that the youngest Homeland speakers in our sample were born, resulting in them having missed out on this change. This highlights that both Homeland and Heritage varieties are dynamic and may develop in different directions. Additionally, this study helps complete the picture previously reported for variation between overt (single or doubled) and null subjects in these two varieties: an ongoing decrease in null subject rates in the Homeland variety and stability in the Heritage variety (Nagy et al. 2018).

1 Introduction

Previous research on Heritage languages in Toronto has shown that for most languages, features do not undergo significant changes when they are transmitted from Homeland to Heritage speakers (Nagy 2014; Kang and Nagy 2016; Nagy et al. 2018). This raises the question of whether this general stability in transfer applies across different language features or whether it is specific to the particular features already investigated, with other unexplored features less resilient in the face of transfer. In order to address this question, we are investigating the use of a yet-unstudied variable in Faetar: Subject doubling, which refers to the phenomenon of two co-referential subject realizations co-occurring in the same clause (De Vogelaer 2010: 222), as in (1).

(1) /la fen i alav pa a laʃkol/
   ‘The woman she did not go to school
   (F2M42A)¹

We investigate subject doubling because Faetar is unusual in that speakers produce more null subjects when the subject represents new information than when it represents old information, violating cross-

¹ Unless otherwise specified, examples are taken from the Heritage Language Variation and Change corpus (Nagy 2011), and contain five pieces of information about the speaker: The Heritage Language of the speaker (here represented as F for Faetar), their generation, their gender, their age, and a unique identifier for otherwise identically-labeled speakers. As there is no standard orthography for Faetar, examples from all sources are transcribed here in IPA for ease of comparison.
linguistic pragmatic expectations (Heap and Nagy 1998; Nagy et al. 2018). The current investigation expands our understanding of variation in the subject context. Subject doubling is an ideal test case for potential differences between Homeland and Heritage speakers since it is possible in Faetar, but very rare in English, the dominant language in Toronto, Canada, where all of the Heritage speakers reside. Because subject doubling is rare in English, we might expect the Heritage speakers, who are in contact with English, to use less doubling in their Faetar than the Homeland speakers.\(^2\)

Previous work on subject doubling in other Romance languages has shown that there are a variety of factors that condition this variable (Kato 1999; Nagy and Blondeau 1999; Auger 2003a; Auger and Villeneuve 2010; \textit{inter alia}), but to date these analyses have focused primarily on Picard and different varieties of French. The present research aims to address this gap in the literature by extending the analysis of doubling to a lesser-studied language. By comparing the Heritage and Homeland communities across communities and generations, we shed light on how variation in Homeland varieties is transmitted. Using quantitative comparative variationist methods, we address the following research questions: (i) what constraints account for the realization of double subjects in Faetar? More specifically, can the existing constraints established in the literature for other Romance languages (such as Picard and French) account for the variability in Faetar double subject realization? (ii) Are there differences in double subject realization between the Heritage and Homeland Faetar speakers? If so, what are they?

The structure of the paper is as follows. In section 2, we provide a brief background of Faetar. Section 3 discusses the variable context in greater detail. Section 4 comprises a literature review of previous studies on subject doubling, and earlier considerations of the constraints under investigation for the present study. Section 5 presents the results of both the distributional and multivariate analyses. Section 6 discusses the results in the context of our original research question and lays out our conclusions.

2 Background: Faetar, an understudied variety of Francoprovençal

Faetar is a variety of Francoprovençal, a language group which has historically been spoken in regions of France, Switzerland and Italy (Kasstan 2016: iii). It is spoken in two small communities in the Italian province of Apulia, Faeto and Celle di San Vito, where it has had contact with southern Italian for some 600 years. Though it has been largely maintained within these communities, thanks in part to the relative geographical isolation provided by the surrounding mountains, there is concern that the language is on the decline. Zulato et al. (2018) point out that Francoprovençal is “endangered everywhere it is spoken” (p. 11); Faetar, in particular, has less than 1,000 native speakers worldwide (Zulato et al. 2018: 18). This makes the need for the documentation and study of Faetar especially acute: while other varieties of Francoprovençal have received substantial attention (see Favre 2002; Diémoz 2009; Bert and Costa 2014; Bichurina 2015; \textit{inter alia}), Faetar leaves much to be explored.

As reported in Nagy (2011: 369), it is often claimed by native speakers that younger generations of Homeland speakers have lost a substantial amount of vocabulary due to the influence of Italian. However, research shows that there are a great number of structural distinctions that survive. One notable difference between Faetar and Italian is that the former, like Picard and French, allows for subject doubling (Marzys 1981: 58; Heap and Nagy 1998: 291), while southern Italian varieties do not (Roberge 1990, Burzio 2012).

3 Data and Method

The data for this project are taken from two corpora of spontaneous speech: one consisting of sociolinguistic interviews with 21 Homeland speakers (collected in Faeto in the 1990s), the other consisting of sociolinguistic interviews with 13 Heritage speakers, including eight first generation speakers and five second generation speakers (collected in Toronto in the 2010s) (Nagy 2011). Together, these materials form part of the Heritage Language Variation and Change (HLVC) corpus, a unique corpus of conversational

\(^2\) A query for pronominal subject doubling in the 450 million-word \textit{Corpus of Contemporary American English} (Davies 2016) corroborates the low rate of this phenomenon in English.
speech in ten Heritage languages (Faetar, Italian, Cantonese, Hungarian, Korean, Russian, Polish, Portuguese, Tagalog and Ukrainian) spoken in Toronto, Ontario (Nagy 2009).

The Heritage Faetar data is separated into two generations: Generation 1 comprises speakers who were born in the homeland, moved to the Greater Toronto Area after age 18, and have lived there for at least 20 years. Generation 2 speakers are those who were either born in the Greater Toronto Area or came from the homeland before age 6, and whose parents qualify as Generation 1 (Nagy 2011). While the Generation 1 and 2 speakers all speak English as well as Faetar, in all cases, they are sufficiently fluent in Faetar to carry a one-hour long conversation in the language. By contrast, the Homeland speakers are individuals born and raised in Faeto, Italy, and had little to no contact with English at the time of data collection. Subject tokens were extracted from 34 speakers, resulting in 976 Homeland and 899 Heritage tokens, respectively, for a total of 1,875 tokens.

To date, there has not been any quantitative research done on subject doubling in Faetar (but cf. Maryzs 1981 for a qualitative account). To fill this gap in the research literature, we explore all possible constraints through a strictly quantitative lens. The data were thus coded for all available social factors (age (continuous), gender (binary, as perceived by the interviewer), and community membership (Homeland vs. Generation 1 vs. Generation 2) as well as linguistic factors previously shown to be significant in subject doubling in other languages (information status, grammatical person, negation, tense, and intervening material). The data were then analyzed using mixed effects logistic regression in Rbrul (Johnson 2009) in order to determine the relative strength of these factors in accounting for the realization of subject doubling.

3.1 Speaker Sample

Table 1 presents an overview of the speaker sample. The Heritage speakers are much older than the Homeland speakers: the average age of the Heritage speakers is 65 (ranging from 32-92), while the average age of the Homeland speakers is 39 (ranging from 11-77). This is because the first waves of Faetar speakers emigrated to Toronto from Italy between the 1950s and 1970s and were, by definition, above the age of 18 when they left (Iannozzi 2016: 2). Their data was collected about 25 years later.

| Community     | Men | Women | Total |
|---------------|-----|-------|-------|
| Homeland      | 10  | 11    | 21    |
| Heritage G1   | 4   | 4     | 8     |
| Heritage G2   | 3   | 2     | 5     |
| Total         | 17  | 17    | 34    |

3.2 The linguistic variable

In Faetar, there are three possible realizations of doubled subjects: noun + weak pronoun (2, repeated from 1), strong + weak pronoun (3), and the demonstrative pronoun + weak pronoun (4). Both the strong and weak forms are able to occur adjacently without emphatic effect (Heap and Nagy 1998: 293). In the analysis, these forms are grouped together as “doubled subjects” and treated as the application value. Generally speaking, strong and weak forms are categorized based on a combination of morphological and

---

3 Previous research has suggested that the use of a strong pronominal subject form versus a weak pronominal subject form has no effect on the meaning of the sentence (Maryzs 1981: 58). We further extrapolate from this that the presence of both a strong form and a weak form in adjacent position would have no effect on the sentence’s general meaning, and do not consider the effects of emphasis.
syntactic characteristics (cf. Cardinaletti and Starke 1994), and in Faetar can be identified by their surface order (strong precedes weak) and a difference in vowel quality (full vowels for strong, schwa for weak), the dropping of either the coda consonant for monosyllabic forms, or the dropping of the second syllable for disyllabic forms.

(2) /la fen i alav pa a laʃkol/  
‘The woman she did not go to school’  
(F2M42A)

(3) /ti ttə mindʒ/  
‘You you eat/are eating’  
(Morosi 1890-92: 64)

(4) /sel i λe me miłɔw da set/  
‘That it is better than this’  
(Nagy 2000: 98)

Examples of single subject realizations in Faetar are single weak pronouns (5), single strong pronouns (6), single demonstrative pronouns (7), or single nouns (or NP) (8). There is also the zero variant (9), which is retained in this analysis as part of the variable context, following previous research (Maryzs 1981). All of the single and zero realizations are grouped together and treated as the non-application value.

(5) /i deʒittə a ssu mari/  
‘She said to her husband’  
(Castielli et. al. 1978: 106)

(6) /jɔˈɾə nuʃ uʃ fa fɨtə/  
‘Now we’ve partied’  
(Castielli et. al. 1978: 99)

(7) /set aetə sel ko dʒə viaw ajər/  
‘This is the one that I saw yesterday’  
(Nagy 2000: 98)

(8) /neom stəvə dʒəfə a me/  
‘Naomi was saying to me’  
(F2F51A)

(9) /Ø mɪndʒ/  
‘Ø eat’  
(Heap and Nagy 1998: 293)

3.3 Exceptional distributions

Generic grammatical person tokens (10) and expletive grammatical person tokens (see 11) were excluded from the analysis as they are not able to be doubled or null and are therefore not part of the variable context. Embedded clauses (12) and future tense tokens (13) were also excluded on the basis of very low numbers of tokens in the data.

(10) /o piówə/  
‘It’s raining’  
(Heap and Nagy 1998: 293)

(11) /alor aijəta pura na bella dʒəɾl/  
‘Then, there is also a pretty pitcher’  
(F0M77A)4

(12) /lu tʃin i vad i award a mandʒ/  
‘The dog, he goes/is waiting to eat’  
(F0M23B)

(13) /dʒə va dʒəf/  
‘I will say’  
(F0F40A)

We also excluded a small number of cases where the pattern of doubling was extremely unusual (e.g. noun + strong pronoun, as in (14), as these are likely to have been speech errors.

4 Examples (14-16) are taken from the Homeland corpus (Nagy 1996).
4 Situating the linguistic variable

In the following subsections, we consider previous analyses of subject doubling and discuss the constraints shown to be significant in the varieties discussed. Subject doubling is not part of Standard English, but has been described in some English dialects (Wolfram and Christian 1976; Southard and Muller 1998). However, the lack of quantitative studies on subject doubling in English prevents us from considering any potential “English origin” constraints. We therefore set aside the specifics of the potential contact effects for the Heritage speakers, and do not provide a comparison of doubling within Toronto English.

4.1 Previous analyses of subject doubling

Previous investigations of subject doubling have focused on a subset of the possible subject doubling cases, constraining the dependent variable in different ways. There is variability, for example, in researchers’ inclusion of null subjects, of first and second person contexts, and of single weak pronouns. Previous studies also vary on whether they include left dislocation.5

Some researchers have focused exclusively on doubling that involves full lexical DPs, excluding pronoun doubling from the envelope of variation (Culbertson 2010). Other researchers excluded lexical DP subject doubling, describing exclusively pronoun doubling (Maryzs 1981; Houze 2016). The majority of investigations, however, have included both subject doubling involving lexical NPs and pronoun doubling (see Nadasdi 1995; Nagy and Blondeau 1999; Nagy et al. 2003; Auger and Villeneuve 2010; Zahler 2014). We follow the latter in including both pronominal and nominal subject realizations.

The inclusion of various grammatical persons has also been dealt with variably in the literature. Several researchers have chosen to exclude first and second person contexts from consideration, focusing exclusively on doubling with third person subjects (Nadasdi 1995; Nagy et al. 2003; Zahler 2014). This decision has been justified by the observation that first and second person pronouns are obligatorily doubled in some dialects of French (Nadasdi 1995), and variation is therefore only expected in the third person. Since this is not the case for Faetar, we include first, second, and third person in the present analysis.

The question of what to include as ‘non-doubled’ contexts has also been dealt with in several ways. Some authors have considered only single preverbal nominal subjects which are able to co-occur with a subject clitic (i.e., lexical NPs + weak pronouns), excluding single weak pronouns that stand on their own (Nadasdi 1995; Nagy et al. 2003; Zahler 2014). Others elected to include single weak pronouns as part of the envelope of variation (Nagy and Blondeau 1999; Houze 2016). Maryzs (1981) goes a step further by including the zero variant as a ‘non-doubled’ context, grouped together with single subject realizations. Since Maryzs’ investigation is the only one to consider Faetar subject doubling, we follow this methodology in retaining single pronoun (both single weak and single strong) realizations, as well as the null variants.

Finally, the question has arisen as to whether left dislocation should be considered as a case of subject doubling or as a separate grammatical structure inactive on subject doubling variation. Left dislocation and subject doubling are similar in terms of their surface structure, but a left-dislocated subject is syntactically analysed as occurring in the topic position, while a true doubled subject is analysed as occupying the subject position (Nadasdi 1995; Auger 2003a, 2003b). There are no universally unambiguous tests to distinguish the two structures, and existing tests rely primarily on prosody (Nagy et al. 2003). Some studies have therefore attempted to exclude left-dislocation (Nadasdi 1995; Auger 2003a; Auger and Villeneuve 2010; Houze 2016), while others did group it in with subject doubling (Nagy and Blondeau 1999; Nagy et al. 2003). We follow the latter group of researchers by including it as part of the variable context.

5 Left dislocation refers to a syntactic structure in which a constituent such as a noun or a pronoun appears to the left of its predicate and has within the same sentence a non-reflexive co-referential pronoun (Duranti and Ochs 1979).
4.2 Independent variables: linguistic factors

The previous literature offers some precedent for what factor groups we might expect to influence subject doubling in Faetar. Due to a lack of quantitative work on subject doubling in Italian, we instead examine the constraints which have been found to significantly influence subject doubling in Picard and French. Specifically, we are focusing on five linguistic constraints: information status, grammatical person, negation, tense, and intervening material.

Information status refers to whether a subject has been mentioned in the discourse. Subjects were coded as being either old or new information. Previous work indicates different directions of effect for this factor, with Barnes (1985) reporting that speakers favour doubling with discourse-new subjects, and Zahler (2014) showing the speakers favour doubling with discourse-old subjects.

Another factor implicated in French subject doubling is grammatical person. Information on this factor is sporadic—there is evidence that some dialects of French undergo categorical subject doubling with first and second person, and on that basis, some research has limited investigation to third person contexts (Nadasdi 1995). Those researchers who did identify grammatical person as significant have found that singular contexts favour subject doubling, while plural contexts disfavour it (Zahler 2014; Houze 2016). There also seems to be evidence that third person contexts favour subject doubling (Nagy and Blondeau 1999). We follow Maryzs (1981) in coding this factor as participant (i.e., 1st and 2nd person) and non-participant (i.e., 3rd person), collapsing singular and (the rarer) plural contexts.

Zahler (2014) found that post-verbal negation and affirmative contexts both favour subject doubling, while pre-verbal negation disfavours it. In contrast to French, Faetar only makes use of post-verbal negation (Nagy 2000). As a consequence, we only distinguish between presence and absence of negation.

Houze (2016) identified verb tense as a significant factor for doubling, but only among younger speakers. Present tense was found to favour subject doubling, while past tense disfavoured it. We retain Houze’s (2016) factor levels, coding tense as present and past (future tense tokens were excluded; see section 3.3).

A final factor implicated in a number of studies is intervening material, i.e. the presence or absence of one or more grammatical items between the subject and the verb. Most studies have found that the presence of an intervening element favours subject doubling, while the absence of an intervening element slightly disfavours it (Nagy and Blondeau 1999; Nagy et al. 2003; Auger and Villeneuve 2010; Zahler 2014). Houze (2016) found the opposite direction of effect in Louisiana French, such that intervening material disfavours subject doubling; however, this effect was only significant for the younger speaker group. For the current analysis, we used the presence or absence of preverbal object pronoun as a proxy for intervening material more generally.

For linguistic factors where the French results have been unambiguous (such as negation and tense), we would expect Faetar subject doubling to follow the above described trends. For factors where the results for French have been less clear (such as for information status, grammatical person, and intervening material) we cannot confidently make predictions about the expected direction of effect for Faetar.

4.3 Independent variables: social factors

The social factors of speaker age and gender were also considered. Given that the results for prodrop showed the same change in progress taking place between the Homeland and the Heritage communities—that is, a decrease in use of null subjects over time (Nagy et al. 2018)—the inclusion of age as a factor enables us to test whether or not subject doubling is also a change in progress and if it is mirrored in both communities. Further, age was found to be a significant factor for the occurrence of doubling in the Saguenay dialect of Québecon French (Auger and Villeneuve 2010), and therefore it may be relevant to account for variation in doubling regardless of its implication in ongoing change. If there is a change in progress, however, gender may prove explanatory: given that women tend to lead in linguistic change (Labov 2001), we would expect the change to be led by women.
As this is the first study to consider subject doubling in Heritage languages, there is little precedent in the literature for the consideration of community or generation membership with respect to this variable. However, other comparative analyses on Heritage and Homeland communities in Toronto has shown that in general, there are few differences between varieties: this includes aforementioned research on Faetar null subjects (Nagy et al. 2018), as well as considerations of pro-drop and Voice Onset Time (VOT) among Heritage and Homeland Cantonese, Russian, Ukrainian and Italian speakers (Nagy 2014). For pro-drop, different generations of Heritage speakers differed little from the monolingual Homeland speakers (Nagy 2014: 325). In their study of Heritage and Homeland participation in a Korean VOT merger, Kang and Nagy (2016) found the same change in progress in both communities, although the younger Heritage speakers were shown to not advance the change as their Seoul counterparts did.

The present analysis allows us to test whether this general stability can be extended to the variable currently under investigation. Using these trends as a baseline for comparison with subject doubling, we expect the Heritage Faetar speakers to double at similar rates – and with similar constraints – as the Homeland speakers do.

5 Results

5.1 Distributional analysis

Table 2 presents the overall distributions for the Homeland and Heritage speakers.

|                     | Homeland (N=976) | Heritage (N=899) |
|---------------------|------------------|------------------|
| Doubled             |                  |                  |
| %                   | N                | %                | N                | %                | N                |
| 24                  | 268              | 76               | 738              | 9                | 80               | 91               | 819              |

The Homeland speakers have a much higher overall rate of doubling than the Heritage speakers: 24% to 9%, respectively. This drop in use of subject doubling from the source to the contact variety was expected and is consistent with the findings for the rate of pro-drop across generations (Nagy et al. 2018). In order to determine the differences between generations, the rates between Generation 1 and Generation 2 are also considered (see Table 3).

|                     | Generation 1 (N=614) | Generation 2 (N=285) |
|---------------------|----------------------|----------------------|
| Doubled             |                      |                      |
| %                   | N                    | %                    | N                    | %                    | N                    |
| 8                   | 48                   | 92                   | 566                  | 11                   | 32                   | 89                   | 253                |

Table 3 shows that the rates for the two heritage groups are similar; though the Generation 2 speakers appear to have a slightly higher rate of doubling than the Generation 1 speakers.

Table 4 presents the distributions in each community according to the linguistic factors information status, grammatical person, negation, tense, and preverbal object.
Table 4. Overall distribution of subject doubling (Homeland and Heritage): Linguistic factors

| Linguistic factors | Homeland (N=976) | Heritage (N =899) |
|--------------------|------------------|------------------|
|                    | % | N (doubled) | Total # of tokens in category | % | N (doubled) | Total # of tokens in category |
| Information status |    |            |                            |    |            |                            |
| New                | 29.1 | 211 | 725 | 8.8 | 52 | 590 |
| Old                | 10.8 | 27  | 251 | 9.1 | 28 | 309 |
| Grammatical person |    |            |                            |    |            |                            |
| Non-participant    | 24.9 | 221 | 886 | 9.1 | 46 | 506 |
| Participant        | 18.9 | 17  | 90  | 8.7 | 34 | 393 |
| Presence of Negation |    |            |                            |    |            |                            |
| Absent             | 25.2 | 232 | 920 | 8.8 | 71 | 809 |
| Present            | 10.7 | 6   | 56  | 10.0 | 9 | 90 |
| Tense              |    |            |                            |    |            |                            |
| Present            | 24.7 | 228 | 922 | 8.9 | 42 | 471 |
| Past               | 18.5 | 10  | 54  | 8.9 | 38 | 428 |
| Preverbal object   |    |            |                            |    |            |                            |
| Absent             | 24.7 | 236 | 954 | 9.0 | 78 | 870 |
| Present            | 9.1  | 2   | 22  | 6.9 | 2  | 29  |

For the Homeland community, speakers have elevated rates of subject doubling in contexts of new information, non-participant (third person) contexts, affirmative contexts, in the present tense, and when there is no preverbal object. We would thus expect these to be the directions of effect for these grammatical constraints on doubling in Homeland Faetar. For the Heritage speakers, the effects of the linguistic factors are less pronounced, although trends other than the negation effect are in the same direction. The Heritage speakers have a slightly higher rate of doubling in negative contexts (where they go in the opposite direction of the Homeland speakers) and in the absence of preverbal objects.

It appears at the outset that there is grammatical conditioning that exists in the Homeland variety that is not as strongly attested in the Heritage variety. However, in the above table, both Heritage generations have been grouped together; to investigate whether the weakening of grammatical constraints has progressed slowly from Generation 1 to Generation 2, the two groups were considered separately. The distributions for each of these communities individually is presented in Table 5.

The distributions show that between the two generations, the rate of doubling increases in different directions for every constraint. While the Generation 1 speakers increase their rate of doubling in new information contexts, the Generation 2 speakers increase their rate in old information contexts; the Generation 1 speakers use more doubling in third person contexts, but the Generation 2 use more doubling in first person contexts, and so on. Fisher’s Exact Tests on each individual constraint for each of the two generations indicate that none of these trends are statistically significant at a level of \( p < .05 \). This could be due to the small token count or an interaction of effects. Alternatively, it is possible that the use of subject doubling for the Heritage speakers is not governed by the same constraints as the use of Homeland speakers.
Table 5. Overall distribution of subject doubling (Generation 1 and Generation 2)

| Linguistic factors | Generation 1 (N=614) | Generation 2 (N=285) |
|--------------------|-----------------------|-----------------------|
|                    | % (doubled)           | Total # of tokens in category | % (doubled) | Total # of tokens in category |
| Information status |          |                      |          |                      |
| New                | 8.0       | 33                   | 10.7     | 19                   |
| Old                | 7.4       | 15                   | 12.1     | 13                   |
| Grammatical person |          |                      |          |                      |
| Non-participant    | 9.0       | 31                   | 9.3      | 15                   |
| Participant        | 6.3       | 17                   | 13.7     | 17                   |
| Presence of Negation |        |                      |          |                      |
| Absent             | 7.4       | 41                   | 11.8     | 30                   |
| Present            | 11.7      | 7                    | 6.7      | 2                    |
| Tense              |          |                      |          |                      |
| Past               | 9.7       | 31                   | 6.4      | 7                    |
| Present            | 5.7       | 17                   | 14.3     | 25                   |
| Preverbal object   |          |                      |          |                      |
| Absent             | 8.0       | 47                   | 11.0     | 31                   |
| Present            | 3.8       | 1                    | 33.3     | 1                    |

Figure 1 and Table 6 present the distributions according to the social constraints age and gender. Due to the great variability in age, it was not feasible to collapse the ages into groups, necessitating individual plotting.

There does not appear to be a clear pattern by age for the Heritage speakers – with the exception of a small handful of outliers, the rate of doubling amongst the Heritage speakers is quite stable. The Homeland speakers, on the other hand, do exhibit an age effect. Young speakers – particularly those under 30 – have elevated rates of doubling compared to speakers older than 30. This is suggestive of a change in progress amongst the Homeland speakers, with the rate of doubling for the community increasing as age decreases. This difference may be attributed to the smaller number of Heritage speakers available for analysis, especially the dearth of younger speakers.
Table 6 presents the rate of doubling by gender in both communities. There are no apparent gender differences in either group.

### Table 6. Distribution of subject doubling by speaker gender

| Gender | Homeland (N = 976) | Heritage (N = 899) |
|--------|--------------------|--------------------|
|        | Social factors     | Social factors     |
|        | %                  | %                  |
|        | N (doubled)        | N (doubled)        |
|        | Total # of tokens in category | Total # of tokens in category |
| Men    | 25.9               | 8.7               |
|        | 118                | 40                |
|        | 456                | 461               |
| Women  | 23.1               | 9.1               |
|        | 120                | 40                |
|        | 520                | 438               |

This indicates that even if we are dealing with a change in progress in Homeland Faetar, as Figure 1 suggests, this change is not led by women, as is frequently the case with changes from inside the community in more well-documented languages like English (Labov 2001: 292). The next step is to test whether any of these results are significant when all factors are considered simultaneously.

### 5.2 Mixed effects logistic regression

In order to determine if the Homeland and Heritage speakers share the same underlying grammar, we ran separate mixed effects logistic regressions for the Homeland and the Heritage speakers. Following standard sociolinguistic practice, we ran separate models for linguistic and social factors before running them together (Tagliamonte 2012: 131); since the same factors came out as significant, we present the results of the models which considered linguistic and social factors simultaneously.
Table 7 shows the results of the step-up/step-down analysis for the Homeland speakers. Two factors were selected as significant: information status and age. This confirms some of the trends we observed in the distributional analysis, namely that speakers favor subject doubling with discourse-new subjects (FW = .64) and that younger speakers are more likely to double their subjects than older speakers (as age increases by one year, the logodds change by -.029). The results further show that other trends observed in the distributional analysis – such as Homeland speakers showing increased subject doubling in non-participant (third person) contexts, affirmative contexts, in the present tense, and when there is no preverbal object – are not statistically significant. The variance for the random effect of speaker is 0.084.

Table 7. Step-up/step-down analysis of subject doubling (Homeland): Linguistic and social factors with age as a continuous predictor and individual as a random effect

| Input: n/a⁶ | Log likelihood: -490.259 | N = 976 |
|-------------|---------------------------|--------|
|             | logodds | FW | % | N  |
| **Fixed effects:** |          |    |   |    |
| **Information status** |          |    |   |    |
| New         | .564    | .64 | 29.1 | 725 |
| Old         | -.564   | .36 | 10.8 | 251 |
| **Range**   | 28      |    |    |    |
| **Grammatical person** |          |    |   |    |
| Non-participant | .02    | [.51] | 24.9 | 886 |
| Participant  | -.02    | [.50] | 18.9 | 90  |
| **Presence of negation** |          |    |   |    |
| Absent      | .333    | [.58] | 25.2 | 920 |
| Present     | -.333   | [.42] | 10.7 | 56  |
| **Tense**   |          |    |   |    |
| Past        | .12     | [.53] | 18.5 | 54  |
| Present     | -.12    | [.47] | 24.7 | 922 |
| **Preverbal object** |          |    |   |    |
| Absent      | .528    | [.63] | 24.7 | 954 |
| Present     | -.528   | [.37] | 9.1  | 22  |
| **Age**     | -.029   | n/a⁷ | n/a | n/a |
| **Gender**  |          |    |   |    |
| Men         | .281    | [.57] | 25.9 | 456 |
| Women       | -.281   | [.43] | 23.1 | 520 |
| **Random effects:** |          |    |   |    |
| **Speaker** |          |    |   |    |
| SD = 0.601  |          |    |   |    |
| N = 21      |          |    |   |    |

⁶ Rbrul does not produce input probabilities when continuous factors are included in the analysis (Daniel Ezra Johnson, p.c., September 21, 2017).

⁷ Rbrul only reports logodds for continuous predictors, not factor weights (Johnson 2009).
We also ran a mixed effects logistic regression for Heritage speakers. However, the small number of tokens and the unbalanced number of speakers of different ages led to convergence issues. Fisher’s Exact Tests on each linguistic factor found no significant trends. Based on the distributional results, there is no evidence that Heritage speakers, like their Homeland counterparts, double subjects more in discourse-new contexts. Furthermore, there is no evidence for age differences in the Heritage group, which suggests that Heritage speakers are not participating in the change towards subject doubling that is happening in the Homeland variety. Another possibility is that we do not have enough representation of different age groups in the Heritage sample.

6 Discussion

We examined a number of linguistic and social constraints that had previously been implicated in subject doubling in Picard and French. Few of these constraints proved relevant for Faetar. Information status is the only linguistic constraint that emerges as significant for the Homeland speakers (doubling being favored for new information). This shows that doubling is indeed grammatically constrained in the source code, and is similar to previous findings for European French (Barnes 1985). This sheds light on the previous findings regarding null vs. overt subjects: subject doubling, and not just subject overtness, is, in fact, more favoured in the new information context than the old. This marked pattern is not evident in the heritage variety.

Age is also significant for the Homeland speakers: as age increases, the likelihood of subject doubling decreases. This change is not mirrored in the Heritage speakers. The results for the Homeland speakers may reflect either a change in progress or an age-grading effect. We suggest that the former interpretation is more plausible. These results can be contextualized within the findings of Nagy, Iannozzi, and Heap (2018): if the use of nulls subjects is decreasing in apparent time, this entails that realized subjects are increasing in apparent time. While doubled subjects are not the only form of subject realization available in Faetar (single realizations also form part of the variable context) doubled subjects are a type of realized subject. There is increasing concern on behalf of the Faetar speech community that the language is losing ground to Italian (Nagy 2011: 368), and so there may be a socially-driven imperative for speakers to increase their use of distinctive features of the language. Null subjects are a characteristically Italian feature, while doubled subjects are a characteristically Faetar feature, therefore Faetar speakers may favour the latter. A surge in the use of this feature may be driven by younger speakers who wish to express their Faetani identity. Nagy et al. (2018: 15) make a similar proposal for the rates of pro-dropping.

There was also a notable overall decrease in the rate of subject doubling between Homeland and Heritage speakers – the Homeland speakers have over twice the rate of doubling than the Heritage speakers. The initial expectation that the same constraints relevant for Homeland Faetar would also turn out to be operational in Heritage Faetar (albeit in a weaker form) has not been supported. The differences between the communities are substantial. Aside from individual, no constraints – social or linguistic – are significant for the Heritage speakers, while two factors are significant for Homeland speakers. This suggests that while the speakers may have acquired the surface form of the variable context (i.e., they know that it is possible to double their subjects in Faetar), they have not acquired its grammatical or social constraints. However, it may be that it is simply harder to see the effects of the constraints for the Heritage speakers because the overall rate of doubling is lower, around 9% compared to 24% in the Homeland (cf. Table 2) and the sample size is perhaps too small to support significant contrasts. Since the Generation 1 speakers were born and raised in Faeto and did not emigrate to Canada until after the age of 18, we can rule out the possibility of incomplete acquisition – speakers must have acquired the constraints during childhood.

As discussed in earlier sections, for most variables, the HLVC project has not found significant differences (in rates or conditioning effects) between Homeland and Heritage speakers (Nagy 2014; Kang and Nagy 2016; Nagy et al. 2018). This makes our findings unusual. But why is this the case? As previously mentioned, a likely culprit is the lowered rate of doubling in the Heritage speakers overall, which makes

---

8 Subject doubling does occur in Italian, but is not prescriptively permitted in the standard variety and is found in northern and central dialects (Brandi and Cordin 1989; D’Alessandro 2009: 6), not the southern dialects near Faeto.
the determination of grammatical constraints challenging. It may also be that pressure from Italian that has led to the decrease of null subjects over time, with speakers wanting to distance themselves from characteristically Italian features. It is plausible that maintaining a Faetani identity is just as important, if not more so, for diasporic communities as it is for speakers in the Homeland. Therefore, the consistency of the change towards more realized subjects in both communities is maintained. Why, then, does the same pattern not hold for doubling? Perhaps the decrease in null subjects is a change that has been underway for far longer than the increase in subject doubling. Recall that the average age of the Homeland speakers was much younger than that of the Heritage group (at the time of recording). Since the youngest speakers in the Homeland data were born roughly around the time that the Generation 1 Heritage speakers emigrated to Canada, it may be that they are not only the leaders of the change, but also the originators. The Generation 1 speakers would have just missed it and were therefore unable to pass it on to Generation 2, accounting for the lack of an age effect in the Toronto community.

Our results reinforce that when considering the transfer of grammatical variation from Homeland to Heritage varieties we cannot expect that all changes will be mirrored by the diasporic community. However, data from Heritage communities can corroborate the starting point of a change in progress taking place in the Homeland. In this case, the facts that the youngest Homeland speakers use more doubling than older Homeland speakers, combined with the absence of an age effect for Heritage speakers who emigrated at the time those younger Homeland speakers were born, stand as a mutually supportive timeline for the increase in subject doubling over time in Faeto. This analysis therefore showcases one of many possible ways that speakers’ departure from the Homeland can lead to differences in the variety spoken by Heritage speakers.

Acknowledgements. We are grateful to the research assistants who have helped transcribe and code the data examined here. They are listed at http://projects.chass.utoronto.ca/ngn/HLVC/3_3_former_ra.php#faetar_fieldworkers. Data collection was supported by SSHRC Standard Research Grant 410-2009-2330 and Salvatori Research Grants (1992, 1993, 1994) to the fifth author. We would also like to thank our anonymous reviewer, David Heap, Michael Iannozzi, Jonathan Kasstan, Gerald Richard, Luke Zhou, and audiences at NWAV 46 in Madison, WI, the University at Buffalo Sociolinguistics Lab, and the Western University Linguistics Speaker Series for helpful comments on earlier versions of this paper.

References

Auger, Julie. 2003a. Le redoublement des sujets en picard. Journal of French Language Studies 13(3): 381–404.

Auger, Julie. 2003b. Les pronoms clitiques sujets en picard: Une analyse au confluent de la phonologie, de la morphologie et de la syntaxe. Journal of French Language Studies 13(1): 1–22.

Auger, Julie, and Anne-José Villeneuve. 2010. La double expression des sujets en français saguenéen: Étude variationniste. In Hétérogénéité et homogénéité dans les pratiques langagières: Mélanges offerts à Denise Deshaies, ed. Wim Remysen and Diane Vincent, 67–86. Québec, QC: Presses de l’Université Laval.

Barnes, Betsy K. 1985. The pragmatics of left-detachment in spoken Standard French. Amsterdam: John Benjamins.

Bert, Michel, and James Costa. 2014. What counts as a linguistic border, for whom, and with what implications? Exploring Occitan and Francoprovençal in Rhône-Alpes, France. In Language, borders and identity, ed. Dominic Watt and Carmen Llamas, 186–204. Edinburgh: Edinburgh University Press.

Bichurina, Natalia. 2015. Le francoprovençal entre la France, la Suisse et l'Italie: Langue diffuse, langue focalisée et enjeux de normalisation. Nouvelles du Centre d'études francoprovençales René Willien 71(1): 7–24.
Brandi, Luciana and Patrizia Cordin. 1989. Two Italian dialects and the null subject parameter. In The null subject parameter, ed. Osvaldo Jaeggli and Kenneth J. Safir, 111–142. Dordrecht: Kluwer.

Burzio, Luigi. 1986. Italian syntax: A government-binding approach. Dordrecht: Reidel.

Cardinaletti, Anna, and Michal Starke. 1994. The typology of structural deficiency: A case study of the three classes of pronouns. In Clitics in the languages of Europe, 8, ed. Henk van Riemsdijk, 145–234. Berlin: De Gruyter.

Castielli, Raffaeli, Michele Melillo, and Armistizio Matteo Melillo. 1978. Storia e cultura dei francoprovenzali di Celle e Faeto: Un saggio storico culturale. Manfredonia: Atlantica.

Culbertson, Jennifer. 2010. Convergent evidence for categorical change in French: From subject clitic to agreement marker. Language 86(1): 85–132.

D’Alessandro, Roberta. 2009. Subject doubling in Italo-Romance varieties. Paper presented at Comparative Syntax of Doubling, Meertens Instituut, Amsterdam.

De Vogelaer, Gunther De. 2010. Does grammaticalization need analogy? Different pathways on the ‘pronoun/agreement marker’-cline. In Grammaticalization: Current views and issues, ed. Katerina Stathi, Elke Gehweiler, and Ekkhard König, 221–240. Amsterdam: John Benjamins.

Diémoz, Federica. 2009. Le pronom personnel sujet de la 3e personne du singulier et le sujet neutre en francoprovençal valaisan: Étude morphosyntaxique. In Langues et cultures de France et d’ailleurs: Hommage à Jean-Baptiste Martin, ed. Claudine Fréchet and Jean-Baptiste Martin, 177–193. Lyon: P.U.L.

Duranti, Alessandro, and Elinor Ochs. 1979. Left-dislocation in Italian conversation. In Discourse and syntax, 377–416. Brill.

Favre, Saverio. 2002. Le francoprovençal valdôtain et l’influence du piémontais. In XIV e XV réscontranterrassionalan sla lenga e la literatura piemontèisa, ed. Gianreno P. Clivio, Dario Pasero, and Censis Pich, 187–196. Ivrea: La Slòira.

Heap, David, and Naomi Nagy. 1998. Subject pronoun variation in Factar and Francoprovençal. In Papers in sociolinguistics. NWAVE-26 à Université Laval, ed. Claude Paradis, Diane Vincent, Denise Deshaies, and Marty Laforest, 291–300. Québec, QC: Nota Bene.

Houze, Beth. 2016. Quantifying linguistic change: Subject pronoun doubling in Louisiana French over time. Masters Forum Paper, University of Toronto, Toronto.

Iannozzi, Michael. 2016. Pro-drop in Factar in Canada: A study of a heritage language in contact. Western Papers in Linguistics 1(2): Article 5. Retrieved from http://ir.lib.uwo.ca/wpl_clw/vol1/iss2/5

Kasstan, Jonathan. 2015. Variation and change in Francoprovençal: A study of an emerging linguistic norm. Doctoral dissertation, University of Kent, Kent.

Johnson, Daniel Ezra. 2009. Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis. Language and Linguistics Compass 3(1): 359–383.

Kang, Yoonjung, and Naomi Nagy. 2016. VOT merger in Heritage Korean in Toronto. Language Variation and Change 28(2): 249–272.

Kato, Mary Aizawa. 1999. Strong and weak pronouns in the null subject parameter. Probus 11(1): 1–38.

Labov, William. 2001. Principles of linguistic change. Volume 2: Social factors. Oxford: Blackwell.

Marzys, Zygmunt. 1981. Les pronoms personnels sujets dans le parler francoprovençal de Faeto et Celle. Vox Romanica 40: 48–58.

Morosi, Giuseppe. 1890–1892. Il dialetto franco-provenzale di Faeto e Celle, nell’Italia meridionale. Archivio Glottologico Italiano XII: 33–75.

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

Nagy, Naomi. 1996. Language contact and language change in the Factar speech community. Doctoral dissertation, University of Pennsylvania, Philadelphia, PA.

Nagy, Naomi. 2000. Factar. Munich: Lincom Europa.
Nagy, Naomi. 2009. Languages we study. Retrieved from http://projects.chass.utoronto.ca/ngn/HLVC/1_2_languages.php
Nagy, Naomi. 2011. Lexical change and language contact: Faetar in Italy and Canada. Journal of Sociolinguistics 15(3): 366–382.
Nagy, Naomi. 2014. A sociolinguistic view of null subjects and VOT in Toronto heritage languages. Lingua 164(B): 309–327.
Nagy, Naomi, and Hélène Blondeau. 1999. Double subject marking in L2 Montreal French. University of Pennsylvania Working Papers in Linguistics 6(2): 93–108.
Nagy, Naomi, Hélène Blondeau, and Julie Auger. 2003. Second language acquisition and “real” French: An investigation of subject doubling in the French of Montreal Anglophones. Language Variation and Change 15(1): 73–103.
Nagy, Naomi, Michael Iannozzi, and David Heap. 2018. Faetar null subjects: A variationist study of a heritage language in contact. In Francoprovençal: documenting a contact variety in Europe and North America. Special issue of International Journal of the Sociology of Language, ed. Jonathan Kasstan and Naomi Nagy, 249: 31–47.
Roberge, Yves. 1990. The syntactic recoverability of null arguments. Kingston & Montreal: McGill-Queen's University Press.
Southard, Bruce, and Al Muller. 1998. Blame it on Twain: Reading American dialects in The Adventures of Huckleberry Finn. In Linguistics at work: A reader of applications, ed. Dallin D. Oaks, 565–573. Harcourt Brace.
Tagliamonte, Sali. 2012. Variationist sociolinguistics: Change, observation, interpretation. Malden, MA: Wiley-Blackwell.
Wolfram, Walt, and Donna Christian. 1976. Appalachian speech. Arlington, VA: Center for Applied Linguistics.
Zahler, Sara. 2014. Variable subject doubling in spoken Parisian French. University of Pennsylvania Working Papers in Linguistics 20(1): 360–371.
Zulato, Alessia, Jonathan Kasstan, and Naomi Nagy. 2018. An overview of Francoprovençal vitality in Europe and North America. International Journal of the Sociology of Language 249: 11–30.